GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 12:03:35 ; Search time 158 Seconds
(without alignments)
447.276 Million cell updates/sec

Title: US-10-030-605A-21
Perfect score: 1108
Sequence: 1 MGRIKFKEDRGFQGHYYSCN.....PNSSSVDKLAAALEHHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_23Sep04:*
1: geneseqp1980s:*
2: geneseqp1990s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1108	100.0	197	4	AAB46850	Aab46850 Bovine ga
2	1066	96.2	197	4	AAB46851	Aab46851 Bovine ga
3	986	89.0	198	4	AAB46848	Aab46848 Bovine ga
4	944	85.2	198	4	AAB46849	Aab46849 Bovine ga
5	769.5	69.4	173	7	ADD47108	Add47108 Rat Prote
6	769.5	69.4	173	7	ADE57551	Ade57551 Rat Prote
7	769.5	69.4	173	7	ADD47114	Add47114 Rat Prote
8	769.5	69.4	173	7	ADD47112	Add47112 Rat Prote
9	769.5	69.4	173	7	ADD47118	Add47118 Rat Prote
10	769.5	69.4	173	7	ADE83335	Ade83335 Rat Prote
11	769.5	69.4	173	7	ADE57555	Ade57555 Rat Prote
12	757.5	68.4	173	7	ADD47104	Add47104 Rat Prote
13	727.5	65.7	174	4	ABG21006	Abg21006 Novel hum
14	727.5	65.7	174	7	ADD47106	Add47106 Human Pro
15	714.5	64.5	362	7	ADC31213	Adc31213 Human nov
16	710.5	64.1	173	7	ADD47110	Add47110 Human Pro
17	710.5	64.1	173	7	ADE83337	Ade83337 Human Pro
18	710.5	64.1	173	7	ADD47116	Add47116 Human Pro
19	710.5	64.1	173	7	ADE57553	Ade57553 Human Pro
20	705	63.6	308	4	ABG21004	Abg21004 Novel hum
21	680	61.4	199	4	ABG21007	Abg21007 Novel hum
22	604	54.5	145	4	ABG21005	Abg21005 Novel hum
23	320	28.9	507	7	ADC31503	Adc31503 Human nov
24	320	28.9	616	7	ADM04410	Adm04410 Human pro
25	320	28.9	781	5	ABB72344	Abb72344 Murine pr
26	320	28.9	888	8	ADO44168	Ado44168 Structura
27	318.5	28.7	420	5	ABP69298	Abp69298 Human pol
28	300	27.1	205	2	AAW17523	Aaw17523 Human bet
29	297	26.8	204	7	ADE63507	Ade63507 Rat Prote
30	297	26.8	205	8	ADP22671	Adp22671 Golden ha
31	297	26.8	211	8	ADP22673	Adp22673 Golden ha
32	288	26.0	252	5	AAU11447	Aau11447 Human cry
33	281.5	25.4	215	2	AAW17522	Aaw17522 Human bet
34	278.5	25.1	542	6	ABR41645	Abr41645 Human DIT
35	278.5	25.1	962	7	ADE29065	Ade29065 Human AIM
36	278.5	25.1	1723	6	ABG74680	Abg74680 Human CGD
37	277	25.0	250	8	ADP22669	Adp22669 Golden ha
38	276.5	25.0	1080	4	AAM00803	Aam00803 Human bon
39	276.5	25.0	1637	4	AAM00916	Aam00916 Human bon
40	276.5	25.0	2263	4	AAM79000	Aam79000 Human pro
41	257	23.2	197	4	AAB47135	Aab47135 CDIFF-16,
42	257	23.2	197	4	ABG08246	Abg08246 Novel hum
43	257	23.2	197	8	ADQ19127	Adq19127 Human sof
44	253	22.8	97	3	AAB58947	Aab58947 Breast an
45	210.5	19.0	511	7	ADB64185	Adb64185 Human pro

ALIGNMENTS

RESULT 1
AAB46850
ID AAB46850 standard; protein; 197 AA.
XX
AC AAB46850;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline mutant protein Mu12A-His fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 18-19; 28pp; German.
XX
CC This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene

CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 100.0%; Score 1108; DB 4; Length 197;
Best Local Similarity 100.0%; Pred. No. 5.7e-110;
Matches 197; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180

Qy        181 SSVDKLAAALEHHHHHH 197
              |||||||||||||||||
Db        181 SSVDKLAAALEHHHHHH 197
```

RESULT 2
AAB46851
ID AAB46851 standard; protein; 197 AA.
XX
AC AAB46851;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline protein WT-HIS fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 19-20; 28pp; German.
XX
CC This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 197 AA;

Query Match 96.2%; Score 1066; DB 4; Length 197;
Best Local Similarity 96.4%; Pred. No. 1.8e-105;
Matches 190; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | |||||||| | |:||||||||||||||||||| ||||||||||||||||||||
Db          1 MGKITFYEDRGFQGHCYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYSDPNS 180

Qy        181 SSVDKLAAALEHHHHHH 197
              |||||||||||||||||
Db        181 SSVDKLAAALEHHHHHH 197
```

RESULT 3
AAB46848
ID AAB46848 standard; protein; 198 AA.
XX
AC AAB46848;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline mutant Mu12A protein fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.
XX
PS Claim 11; Page 16-17; 28pp; German.

XX
CC This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 198 AA;

Query Match 89.0%; Score 986; DB 4; Length 198;
Best Local Similarity 99.4%; Pred. No. 6.5e-97;
Matches 175; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          6 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 65

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 125

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYS 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        126 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYA 181
```

RESULT 4
AAB46849
ID AAB46849 standard; protein; 198 AA.
XX
AC AAB46849;
XX
DT 26-APR-2001 (first entry)
XX
DE Bovine gamma-crystalline protein fragment.
XX
KW Gamma-crystalline; mutant; beta-leaflet; cosmetic; bioseparation;
KW biosensor; pollution detection; pollution control; gene therapy;
KW intracellular immunization.
XX
OS Bos taurus.
XX
PN DE19932688-A1.
XX
PD 18-JAN-2001.
XX
PF 13-JUL-1999; 99DE-01032688.
XX
PR 13-JUL-1999; 99DE-01032688.
XX
PA (FIED/) FIEDLER U.
PA (RUDO/) RUDOLPH R.
XX
PI Rudolph R, Fiedler U, Boehm G, Reimann C;
XX
DR WPI; 2001-148304/16.
XX
PT Mutated proteins having beta-leaflet structure and related nucleic acid,
PT have new or improved properties, e.g. antibody-like specific binding or
PT catalytic activity.

XX
PS Claim 11; Page 17-18; 28pp; German.
XX
CC This invention describes a novel protein (I) with beta-'leaflet'
CC structure having surface-exposed amino acids, present in at least two
CC surface-exposed beta-strands of a surface-exposed beta-leaflet. The
CC protein is altered by targeted mutagenesis so that it has new, or
CC improved, specific binding, catalytic or fluorescent properties. The
CC invention also describes (1) DNA (II) that encodes (I); (2) RNA (III)
CC derived from (II); (3) prokaryotic and eukaryotic vectors and cells that
CC contain (II) or (III), or their fragments that encode a functional region
CC of (I); and (4) method for producing (I). (I) are useful for diagnosis
CC and therapy, in cosmetics, bioseparation and biosensors, and for
CC pollution detection and control, e.g. for specific targeting of gene
CC therapy vectors and for intracellular immunization. (I) can be provided
CC with new or improved specific antibody-like binding, catalytic or
CC fluorescent properties, without the cost and difficulties associated with
CC producing complete or recombinant antibodies. (I) are relatively small
CC (20 kDa) and can be expressed with other components as multifunctional
CC fusions. They have good stability against low pH, denaturing agents and
CC high temperatures, conditions under which antibodies are unstable
XX
SQ Sequence 198 AA;

Query Match 85.2%; Score 944; DB 4; Length 198;
Best Local Similarity 95.5%; Pred. No. 2e-92;
Matches 168; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | |||||||| | |:||||||||||||||||| |||||||||||||||||||
Db          6 MGKITFYEDRGFQGHCYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 65

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         66 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 125

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYS 176
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||:
Db        126 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFYA 181
```

RESULT 5
ADD47108
ID ADD47108 standard; protein; 173 AA.
XX
AC ADD47108;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P10065, SEQ ID NO 12798.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P10065.

XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 173 AA;

Query Match 69.4%; Score 769.5; DB 7; Length 173;
Best Local Similarity 77.6%; Pred. No. 7.9e-74;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

```
Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:|||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::||||||||||||:|||||||||||:|||| |||||:|||||||||| |
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173
```

RESULT 6
ADE57551
ID ADE57551 standard; protein; 173 AA.
XX
AC ADE57551;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein P10065, SEQ ID NO 3413.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; P10065.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 173 AA;

Query Match 69.4%; Score 769.5; DB 7; Length 173;
Best Local Similarity 77.6%; Pred. No. 7.9e-74;
Matches 135; Conservative 20; Mismatches 18; Indels 1; Gaps 1;

```
Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:|||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::||||||||||||:|||||||||||:|||| |||||:|||||||||| |
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173
```

RESULT 7
ADD47114
ID ADD47114 standard; protein; 173 AA.
XX
AC ADD47114;
XX
DT 29-JAN-2004 (first entry)
XX

```
DE   Rat Protein P10065, SEQ ID NO 12804.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P10065.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 173 AA;

  Query Match             69.4%;  Score 769.5;  DB 7;  Length 173;
  Best Local Similarity   77.6%;  Pred. No. 7.9e-74;
  Matches  135;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:|||||| |||||||||| |||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:||||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::|||||||||||:||||||||||:|||| |||||:|||||||||||| |
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173

RESULT 8
ADD47112
ID   ADD47112 standard; protein; 173 AA.
XX
AC   ADD47112;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein AAA40981, SEQ ID NO 12802.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAA40981.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 173 AA;

  Query Match             69.4%;  Score 769.5;  DB 7;  Length 173;
```

```
  Best Local Similarity   77.6%;  Pred. No. 7.9e-74;
  Matches  135;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              ||||||||||:||||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::|||||||||||:||||||||||:|||| |||||:||||||||||| |
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173


RESULT 9
ADD47118
ID   ADD47118 standard; protein; 173 AA.
XX
AC   ADD47118;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein AAA40981, SEQ ID NO 12808.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAA40981.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 173 AA;

  Query Match             69.4%;  Score 769.5;  DB 7;  Length 173;
  Best Local Similarity   77.6%;  Pred. No. 7.9e-74;
  Matches  135;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              ||||||||||:||||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::|||||||||||:||||||||||:|||| |||||:||||||||||| |
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173


RESULT 10
ADE83335
ID   ADE83335 standard; protein; 173 AA.
XX
AC   ADE83335;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein P10065, SEQ ID NO 10928.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; P10065.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
```

```
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 173 AA;

  Query Match             69.4%;  Score 769.5;  DB 7;  Length 173;
  Best Local Similarity   77.6%;  Pred. No. 7.9e-74;
  Matches  135;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:|||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::|||||||||||||:|||||||||||:|||| |||||:|||||||||||
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173


RESULT 11
ADE57555
ID   ADE57555 standard; protein; 173 AA.
XX
AC   ADE57555;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein AAA40981, SEQ ID NO 3417.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
PR   14-AUG-2001; 2001US-0312147P.
PR   01-NOV-2001; 2001US-0346382P.
PR   26-NOV-2001; 2001US-0333347P.
XX
PA   (GEHO ) GEN HOSPITAL CORP.
PA   (FARB ) BAYER AG.
XX
PI   Woolf C,  D'urso D,  Befort K,  Costigan M;
XX
DR   WPI; 2003-268312/26.
DR   GENBANK; AAA40981.
XX
PT   New composition comprising two or more isolated polypeptides, useful for
PT   preparing a medicament for treating pain in an animal.
XX
PS   Claim 1; Page; 1017pp; English.
XX
CC   The invention discloses a composition comprising two or more isolated rat
CC   or human polynucleotides or a polynucleotide which represents a fragment,
CC   derivative or allelic variation of the nucleic acid sequence. Also
CC   claimed are a vector comprising the novel polynucleotide, a host cell
CC   comprising the vector, a method for identifying a nucleotide sequence
CC   which is differentially regulated in an animal subjected to pain and a
CC   kit to perform the method, an array, a method for identifying an agent
CC   that increases or decreases the expression of the polynucleotide sequence
CC   that is differentially expressed in neuronal tissue of a first animal
CC   subjected to pain, a method for identifying a compound which regulates
CC   the expression of a polynucleotide sequence which is differentially
CC   expressed in an animal subjected to pain, a method for identifying a
CC   compound that regulates the activity of one or more of the
CC   polynucleotides, a method for producing a pharmaceutical composition, a
CC   method for identifying a compound or small molecule that regulates the
CC   activity in an animal of one or more of the polypeptides given in the
CC   specification, a method for identifying a compound useful in treating
CC   pain and a pharmaceutical composition comprising the one or more
CC   polypeptides or their antibodies. The polynucleotide or the compound that
CC   modulates its activity is useful for preparing a medicament for treating
CC   pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC   injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC   therapy). The sequence presented is a rat protein (shown in Table 2 of
CC   the specification) which is differentially expressed during pain. Note:
CC   The sequence data for this patent did not form part of the printed
CC   specification, but was obtained in electronic form directly from WIPO at
CC   ftp.wipo.int/pub/published_pct_sequences.
XX
SQ   Sequence 173 AA;

  Query Match             69.4%;  Score 769.5;  DB 7;  Length 173;
  Best Local Similarity   77.6%;  Pred. No. 7.9e-74;
  Matches  135;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:|||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::|||||||||||||:|||||||||||:|||| |||||:|||||||||||
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173


RESULT 12
ADD47104
ID   ADD47104 standard; protein; 173 AA.
XX
AC   ADD47104;
XX
DT   29-JAN-2004  (first entry)
XX
DE   Rat Protein AAA40988, SEQ ID NO 12794.
XX
KW   Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW   chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
XX
OS   Rattus norvegicus.
XX
PN   WO2003016475-A2.
XX
PD   27-FEB-2003.
XX
PF   14-AUG-2002; 2002WO-US025765.
XX
```

PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; AAA40988.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 173 AA;

Query Match 68.4%; Score 757.5; DB 7; Length 173;
Best Local Similarity 75.3%; Pred. No. 1.5e-72;
Matches 131; Conservative 22; Mismatches 20; Indels 1; Gaps 1;

```
Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |::|  |||||||||||:|| ||||||||:||: | |||||||
Db          1 GKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:||:||| ||| |: : |:|||||:|:|||| ||||||| |||||| ::
Db         61 DYPDYQQWMGFSDSVRSCHLIP-HSSSHRIRIYEREDYRGQMVEITDDCPHLQDRFHFSD 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              || :|:|| |||||||:|||||||||| ||||| |||||||:||||||:||:|
Db        120 FHSFHVIEGYWVLYEMPNYRGRQYLLRPREYRRYHDWGAMNARVGSLRRIMDYY 173
```

RESULT 13
ABG21006
ID ABG21006 standard; protein; 174 AA.
XX
AC ABG21006;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20997.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI; 2001-639362/73.
DR N-PSDB; AAS85193.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 51365; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 174 AA;

Query Match 65.7%; Score 727.5; DB 4; Length 174;
Best Local Similarity 73.6%; Pred. No. 2.5e-69;
Matches 128; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:|   ||||||| :| |:|| |||||| ||||| || |||||||:||| | ||||||
Db          1 MGKITLYEDRGFQGRHYECSSDHPNLQPYLSRCNSARVDSGCWMLYEQPNYSGLQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||| |:||||| :||:|||||| |:|: |:|:|||:|:|||| | |:||  |||||
Db         61 GDYADHQQWMGLSDSVRSCRLIP-HSGSHRIRLYEREDYRGQMIEFTEDCSCLQDRFRFN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDF 174
              |:|||||||||||||: :|||||||| ||:|||| |||| ||:||||||:||
Db        120 EIHSLNVLEGSWVLYELSNYRGRQYLLMPGDYRRYQDWGATNARVGSLRRVIDF 173
```

RESULT 14
ADD47106
ID ADD47106 standard; protein; 174 AA.

XX
AC ADD47106;
XX
DT 29-JAN-2004 (first entry)
XX
DE Human Protein NP_008822, SEQ ID NO 12796.
XX
KW Human; pain; neuronal tissue; gene therapy;
KW spinal segmental nerve injury; chronic constriction injury; CCI;
KW spared nerve injury; SNI; Chung.
XX
OS Homo sapiens.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.
PR 26-NOV-2001; 2001US-0333347P.
XX
PA (GEHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; NP_008822.
XX
PT New composition comprising two or more isolated polypeptides, useful for
PT preparing a medicament for treating pain in an animal.
XX
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a human protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 174 AA;

Query Match 65.7%; Score 727.5; DB 7; Length 174;
Best Local Similarity 73.6%; Pred. No. 2.5e-69;
Matches 128; Conservative 20; Mismatches 25; Indels 1; Gaps 1;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:|   ||||||| :| |:|| |||||| ||||| || |||||||||:||| | ||||||
Db          1 MGKITLYEDRGFQGRHYECSSDHPNLQPYLSRCNSARVDSGCWMLYEQPNYSGLQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||| |:||||| :||:||||||| |:|: |:|:|||:|:|||| | |:||  |||||
Db         61 GDYADHQQWMGLSDSVRSCRLIP-HSGSHRIRLYEREDYRGQMIEFTEDCSCLQDRFRFN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDF 174
              |:|||||||||||||: :|||||||| ||:|||| |||| ||:|||||||:||
Db        120 EIHSLNVLEGSWVLYELSNYRGRQYLLMPGDYRRYQDWGATNARVGSLRRVIDF 173
```

RESULT 15
ADC31213
ID ADC31213 standard; protein; 362 AA.
XX
AC ADC31213;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel polypeptide sequence, SEQ ID NO:1295.
XX
KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnerary;
KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 18.
XX
OS Homo sapiens.
XX
PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX
DR WPI; 2003-371981/35.
DR N-PSDB; ADC30242.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
PT cancer.
XX
PS Claim 20; SEQ ID NO 1295; 1185pp; English.
XX
CC The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889)and the polypeptides they encode (ADC30890-ADC31860). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of peventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the

CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 362 AA;

Query Match 64.5%; Score 714.5; DB 7; Length 362;
Best Local Similarity 72.3%; Pred. No. 1.7e-67;
Matches 125; Conservative 21; Mismatches 26; Indels 1; Gaps 1;

```
Qy          3 RIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGD 62
              :| | ||| |||  |:| ||||||: |||||||||| ||||||||||||||||||||||
Db        191 KITFYEDRDFQGRCYNCISDCPNLRVYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRRGK 250

Qy         63 YPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTEV 122
              ||||| ||| :||::|||:|| || : ::|:|||||:|| |||:|||| : : | | |:
Db        251 YPDYQHWMGLSDSVQSCRIIP-HTSSHKLRLYERDDYRGLMSELTDDCACVPELFRLPEI 309

Qy        123 HSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :||:|||| |||||||:||||||||||:|||| |||  :||||||||| | |
Db        310 YSLHVLEGCWVLYEMPNYRGRQYLLRPGDYRRYHDWGGADAKVGSLRRVTDLY 362
```

Search completed: January 14, 2005, 12:06:24
Job time : 161 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 12:03:35 ; Search time 16 Seconds
(without alignments)
1184.669 Million cell updates/sec

Title: US-10-030-605A-21
Perfect score: 1108
Sequence: 1 MGRIKFKEDRGFQGHYYSCN.....PNSSSVDKLAAALEHHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	943	85.1	175	1	CYBOG	gamma-crystallin I
2	881	79.5	175	2	B24060	gamma-crystallin 1
3	864	78.0	175	1	CYMSG3	gamma-B-crystallin
4	830	74.9	175	1	CYHUG1	gamma-crystallin 1
5	783.5	70.7	174	2	D24060	gamma-crystallin 2
6	778.5	70.3	174	1	CYMSG2	gamma-crystallin 2
7	778.5	70.3	174	2	I56381	gammma-crystallin
8	776.5	70.1	174	2	I83432	gamma-C-crystallin
9	775.5	70.0	174	2	C24060	gamma-crystallin 2
10	774.5	69.9	174	2	A24060	gamma-crystallin 1
11	773.5	69.8	173	1	CYRTG1	gamma-crystallin 1
12	770.5	69.5	173	2	S04266	gamma-crystallin I
13	770.5	69.5	174	1	CYMSG1	gamma-crystallin 1
14	767.5	69.3	174	2	JS0596	gamma-B-crystallin
15	765.5	69.1	174	1	CYMSG4	gamma-crystallin 4
16	764.5	69.0	174	2	S26811	gamma-E-crystallin
17	762.5	68.8	174	2	E24060	gamma-crystallin 4
18	756.5	68.3	174	2	I45881	gamma-C-crystallin
19	752.5	67.9	174	2	S33526	gamma-crystallin -
20	740.5	66.8	170	1	CYRTG2	gamma-crystallin 2
21	726.5	65.6	174	1	CYHUG2	gamma-crystallin 2
22	723.5	65.3	174	2	I77413	gamma-D-crystallin
23	715.5	64.6	174	2	A26912	gamma-crystallin 5
24	713.5	64.4	157	2	B29655	gamma-crystallin I
25	686	61.9	176	2	S45015	gamma-crystallin M
26	684.5	61.8	159	2	I49618	gamma-F-crystallin
27	682.5	61.6	151	2	I49614	gamma-C-crystallin
28	656	59.2	173	2	S45016	gamma-crystallin S
29	645	58.2	175	2	JN0682	gamma3-crystallin
30	639	57.7	168	1	CYFGG2	gamma-crystallin I
31	639	57.7	169	2	T01783	gamma-2-crystallin
32	631	56.9	175	2	JN0681	gamma2-crystallin
33	627	56.6	177	2	S52842	gamma-M1-1 crystal
34	621	56.0	175	2	JN0680	gamma1-crystallin
35	616	55.6	177	2	S52843	gamma-M1-2 crystal
36	613	55.3	172	2	PN0545	gamma4-crystallin
37	610	55.1	173	2	S45017	gamma-crystallin S
38	561	50.6	133	1	CYFGG	gamma-crystallin I
39	554.5	50.0	174	2	S07146	gamma-s-crystallin
40	545.5	49.2	174	2	S18460	gamma-crystallin m
41	543.5	49.1	177	2	JC2357	gamma-crystallin M
42	538	48.6	173	1	CYCAG2	gamma-crystallin m
43	537.5	48.5	178	1	CYBOS	gamma-s-crystallin
44	532.5	48.1	178	2	I50601	gamma-crystallin M
45	531.5	48.0	175	2	JC2354	gamma-crystallin M

ALIGNMENTS

RESULT 1
CYBOG
gamma-crystallin II - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 28-Feb-1986 #sequence_revision 03-Feb-1994 #text_change 09-Jul-2004
C;Accession: A29655; A02928; A90263; A60890; A60815; A02929; S04265
R;Hay, R.E.; Woods, W.D.; Church, R.L.; Petrash, J.M.
Biochem. Biophys. Res. Commun. 146, 332-338, 1987
A;Title: cDNA clones encoding bovine gamma-crystallins.
A;Reference number: A90133; MUID:87270760; PMID:3606621
A;Accession: A29655
A;Molecule type: mRNA
A;Residues: 1-175 <HAY>
A;Cross-references: UNIPROT:P02526; GB:M16894; NID:g162918; PIDN:AAA30476.1; PID:g162919
R;Bhat, S.P.; Spector, A.
DNA 3, 287-295, 1984
A;Title: Complete nucleotide sequence of a cDNA derived from calf lens gamma-crystallin
A;Reference number: A02928; MUID:85026663; PMID:6092016
A;Accession: A02928
A;Molecule type: mRNA
A;Residues: 2-119,'S',121-175 <BHA>
A;Cross-references: GB:K02112; GB:X01036; NID:g162916; PIDN:AAA30475.1; PID:g162917
A;Note: initiator Met not shown
R;Croft, L.R.
Biochem. J. 128, 961-970, 1972
A;Title: The amino acid sequence of gamma-crystallin (fraction II) from calf lens.
A;Reference number: A90263; MUID:73054483; PMID:4674126
A;Accession: A90263
A;Molecule type: protein
A;Residues: 2-17,'Q',19-21,'NN',23,'LQP',28-39,'VHSL',45-46,'MLQ',48-49,'D',51,53-54,56,
A;Note: portions of this sequence were assigned by composition rather than by direct seq
R;Chiou, S.H.; Azari, P.; Himmel, M.E.
J. Protein Chem. 7, 67-80, 1988
A;Title: Physicochemical characterization of gamma-crystallins from bovine lens-hydrodyn
A;Reference number: A60890; MUID:89351571; PMID:3255364
A;Accession: A60890
A;Molecule type: protein
A;Residues: 2-26 <CHI>
R;McDermott, M.J.; Gawinowicz-Kolks, M.A.; Chiesa, R.; Spector, A.
Arch. Biochem. Biophys. 262, 609-619, 1988
A;Title: The disulfide content of calf gamma-crystallin.
A;Reference number: A60815; MUID:88208422; PMID:3364984
A;Accession: A60815
A;Molecule type: protein
A;Residues: 2-26 <MCD>
R;Blundell, T.; Lindley, P.; Miller, L.; Moss, D.; Slingsby, C.; Tickle, I.; Turnell, B.
Nature 289, 771-777, 1981
A;Title: The molecular structure and stability of the eye lens: x-ray analysis of gamma-
A;Reference number: A93247; MUID:81123111; PMID:7464942
A;Contents: annotation; X-ray crystallography, 2.6 angstroms
R;White, H.E.; Driessen, H.P.C.; Slingsby, C.; Moss, D.S.; Lindley, P.F.
J. Mol. Biol. 207, 217-235, 1989

A;Title: Packing interactions in the eye-lens. Structural analysis, internal symmetry an
A;Reference number: S04265; MUID:89293855; PMID:2738925
A;Contents: annotation
C;Comment: The protein has a two-domain beta-structure, folded into four very similar Gr
C;Comment: Evidence from reference A90263 suggests an absence of disulfide bonds. Evider
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;89-129/Domain: crystallin repeat <GK3>
F;130-169/Domain: crystallin repeat <GK4>

Query Match 85.1%; Score 943; DB 1; Length 175;
Best Local Similarity 96.0%; Pred. No. 8.8e-82;
Matches 168; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | |||||||| | |:|||||||||||||||||| ||||||||||||||||||||
Db          1 MGKITFYEDRGFQGHCYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
```

RESULT 2
B24060
gamma-crystallin 1-2 - rat
N;Alternate names: gamma-B-crystallin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: B24060; I83431
R;den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A;Reference number: A92927; MUID:87060933; PMID:3783678
A;Accession: B24060
A;Molecule type: DNA
A;Residues: 1-175 <DEN>
A;Cross-references: UNIPROT:P10066; GB:M19359; NID:g203626; PIDN:AAA40982.1; PID:g203628
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: I60312; MUID:89378747; PMID:2777080
A;Accession: I83431
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-175 <RES>
A;Cross-references: GB:M19359; NID:g203626; PIDN:AAA40982.1; PID:g203628
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

Query Match 79.5%; Score 881; DB 2; Length 175;
Best Local Similarity 88.6%; Pred. No. 6.3e-76;
Matches 155; Conservative 9; Mismatches 11; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| | |:||||||| |||||||:|| ||||||||||||||||||||
Db          1 MGKITFFEDRGFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:|||||||||||:||:|||||||||||||||||||| ||||||||:
Db         61 GDYPDYQQWMGFSDSIRSCRLIPQHSGTYRMRIYERDDFRGQMSEITDDCLSLQDRFHLS 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:|||||:|| |||||||||||||||||||||||||| |||||| |||||||
Db        121 EIHSLNVMEGCWVLYEMPSYRGRQYLLRPGEYRRYLDWGAANAKVGSFRRVMDFY 175
```

RESULT 3
CYMSG3
gamma-B-crystallin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 17-Mar-1987 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: I48359; A02933; S33523
R;Graw, J.; Liebstein, A.; Pietrowski, D.; Schmitt-John, T.; Werner, T.
Gene 136, 145-156, 1993
A;Title: Genomic sequences of murine gamma B- and gamma C-crystallin-encoding genes: pron
A;Reference number: I48359; MUID:94123992; PMID:8293998
A;Accession: I48359
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-175 <RES>
A;Cross-references: UNIPROT:P04344; EMBL:Z22573; NID:g311633; PIDN:CAA80296.1; PID:g3116
R;Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Tsui,
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A;Title: gamma-crystallin family of the mous lens: structural and evolutionary relationsh
A;Reference number: A94021; MUID:85088487; PMID:6096855
A;Accession: A02933
A;Molecule type: mRNA
A;Residues: 41-150,'M',152-175 <BRE>
C;Comment: There are at least seven different gamma crystallins identified in mouse lens
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;89-129/Domain: crystallin repeat <GK3>
F;130-169/Domain: crystallin repeat <GK4>

Query Match 78.0%; Score 864; DB 1; Length 175;
Best Local Similarity 86.3%; Pred. No. 2.5e-74;
Matches 151; Conservative 11; Mismatches 13; Indels 0; Gaps 0;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||| ||| | |:|||||||| |||||||:|| ||||||||||||||||||||
Db          1 MGKITFFEDRSFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |:||||||||:|||||||||||:||:|||||||||||||||||| ||||||| :
Db         61 GEYPDYQQWMGFSDSIRSCRLIPQHSGTYRMRIYEKDDFRGQMSEITDDCLSLQDRFHFS 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:|||||:|| |||||||||||||||||||||||||| |||||| |||||||
Db        121 EIHSLNVMEGCWVLYEMPSYRGRQYLLRPGEYRRYLDWGAANAKVGSFRRVMDFY 175
```

RESULT 4
CYHUG1
gamma-crystallin 1-2 - human
C;Species: Homo sapiens (man)
C;Date: 30-Jun-1988 #sequence_revision 30-Jun-1988 #text_change 09-Jul-2004
C;Accession: A24520
R;den Dunnen, J.T.; Moormann, R.J.M.; Cremers, F.P.M.; Schoenmakers, J.G.G.
Gene 38, 197-204, 1985
A;Title: Two human gamma-crystallin genes are linked and riddled with Alu-repeats.
A;Reference number: A91536; MUID:86056977; PMID:4065573
A;Accession: A24520
A;Molecule type: DNA
A;Residues: 1-175 <DEN>
A;Cross-references: UNIPROT:P07316; GB:M11971; NID:g181114; PIDN:AAA52113.1; PID:g181118
C;Genetics:
A;Gene: GDB:CRYGA; CRYG1
A;Cross-references: GDB:119076; OMIM:123660
A;Map position: 2q33-2q35
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-175/Product: gamma-crystallin 1-2 #status predicted <MPT>

```
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;89-129/Domain: crystallin repeat <GK3>
F;130-169/Domain: crystallin repeat <GK4>

  Query Match             74.9%;  Score 830;  DB 1;  Length 175;
  Best Local Similarity   81.7%;  Pred. No. 4.2e-71;
  Matches  143;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||| |||  | | :||||||||||||||||| |||||:|||||||||||||||
Db          1 MGKITFYEDRAFQGRSYECTTDCPNLQPYFSRCNSIRVESGCWMIYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |:|||||||| :|||||| ||| |:| :||:||:||: ||||||:|||| |:|||||||
Db         61 GEYPDYQQWMGLSDSIRSCCLIPPHSGAYRMKIYDRDELRGQMSELTDDCLSVQDRFHLT 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:||||||||||:|||||:||||||||||||||:||||| |||||||||||| |
Db        121 EIHSLNVLEGSWILYEMPNYRGRQYLLRPGEYRRFLDWGAPNAKVGSLRRVMDLY 175


RESULT 5
D24060
gamma-crystallin 2-2 - rat
N;Alternate names: gamma-D-crystallin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: D24060; I83433; S60744
R;den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A;Reference number: A92927; MUID:87060933; PMID:3783678
A;Accession: D24060
A;Molecule type: DNA
A;Residues: 1-174 <DEN>
A;Cross-references: UNIPROT:P10067; GB:M19359; NID:g203626; PIDN:AAA40984.1; PID:g203630
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: I60312; MUID:89378747; PMID:2777080
A;Accession: I83433
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-174 <RES>
A;Cross-references: GB:M19359; NID:g203626; PIDN:AAA40984.1; PID:g203630
R;Ooki, K.; Amuro, N.; Shimizu, Y.; Okazaki, T.
Biochimie 76, 398-403, 1994
A;Title: High level expression of rat gamma-D-crystallin in Escherichia coli.
A;Reference number: S60744; MUID:95151850; PMID:7849105
A;Accession: S60744
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-174 <OOK>
A;Cross-references: EMBL:X57169; NID:g929827; PIDN:CAA40458.1; PID:g929828
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

  Query Match             70.7%;  Score 783.5;  DB 2;  Length 174;
  Best Local Similarity   78.3%;  Pred. No. 1e-66;
  Matches  137;  Conservative   19;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  |||||||||||:|| |||||||:||: | ||||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:|||||||  | |: |:|:|||:|:|||| | |:||||||||||
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HAGSHRIRLYEREDYRGQMVEFTEDCPSLQDRFHFN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |::||||||| |||||| :|||||||||||||||| |||||||:||||||||||
Db        120 EIYSLNVLEGCWVLYEMTNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRVMDFY 174


RESULT 6
CYMSG2
gamma-crystallin 2 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C;Accession: A02931
R;Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Tsui
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A;Title: gamma-crystallin family of the mous lens: structural and evolutionary relationsh
A;Reference number: A94021; MUID:85088487; PMID:6096855
A;Accession: A02931
A;Molecule type: mRNA
A;Residues: 1-174 <BRE>
A;Cross-references: UNIPROT:P02528; GB:K02584; NID:g192786; PIDN:AAA03228.1; PID:g309198
C;Comment: There are at least seven different gamma crystallins identified in mouse lens
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;88-128/Domain: crystallin repeat <GK3>
F;129-168/Domain: crystallin repeat <GK4>

  Query Match             70.3%;  Score 778.5;  DB 1;  Length 174;
  Best Local Similarity   77.1%;  Pred. No. 3e-66;
  Matches  135;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  |||||||||||:|| |||||||:||: | ||||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:|||||||  | |: |:|:|||:|:|||| | |:||||||||||
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HSSSHRIRIYEREDYRGQMVEITDDCPHLQDRFHFS 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              : || :|:|| ||||||||:|||||||||||||||| |||||||:|||||:||||
Db        120 DFHSFHVMEGYWVLYEMPNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRIMDFY 174


RESULT 7
I56381
gammma-crystallin - rat
N;Alternate names: gamma-E-crystallin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C;Accession: I56381; I83434
R;Moormann, R.J.; den Dunnen, J.T.; Mulleners, L.; Andreoli, P.; Bloemendal, H.; Schoenma
J. Mol. Biol. 171, 353-368, 1983
A;Title: Strict co-linearity of genetic and protein folding domains in an intragenically
A;Reference number: I56381; MUID:84114867; PMID:6319707
A;Accession: I56381
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-174 <RES>
A;Cross-references: UNIPROT:P02528; EMBL:X00271; NID:g56193; PIDN:CAA25073.1; PID:g29591
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: I60312; MUID:89378747; PMID:2777080
A;Accession: I83434
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-174 <RE2>
A;Cross-references: GB:M19359; NID:g203626; PIDN:AAA40985.1; PID:g203631
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication
```

```
  Query Match           70.3%;  Score 778.5;  DB 2;  Length 174;
  Best Local Similarity  77.1%;  Pred. No. 3e-66;
  Matches  135;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  ||||||||||:|| |||||||:||: | ||||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:||||||| |: : |:|||||:|:|||| ||||||| |||||| :
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HSSSHRIRIYEREDYRGQMVEITDDCPHLQDRFHFS 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              : || :|:|| |||||||:||||||||||||||| |||||||:||||||:||||
Db        120 DFHSFHVMEGYWVLYEMPNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRIMDFY 174


RESULT 8
I83432
gamma-C-crystallin - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 09-Jul-2004
C;Accession: I83432
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: I60312; MUID:89378747; PMID:2777080
A;Accession: I83432
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-174 <RES>
A;Cross-references: UNIPROT:P02529; GB:M19359; NID:g203626; PIDN:AAA40983.1; PID:g203629
C;Genetics:
A;Gene: CRY-gamma-C
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

  Query Match           70.1%;  Score 776.5;  DB 2;  Length 174;
  Best Local Similarity  78.3%;  Pred. No. 4.7e-66;
  Matches  137;  Conservative   18;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| | |:|||||| |||||||||| ||||||||||||||||||||
Db          1 MGKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:|||||||||| |||: |||:||::| :| | |:::|| :||||||:
Db         61 GDYPDYQQWMGFSDSIRSCRLIP-HTGSHRMRLYEKEDHKGVMMELSEDCSCIQDRFHLS 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              || ||:|||| |||||||:|||||||||| ||||| ||||::|| ||||||:| |
Db        120 EVRSLHVLEGCWVLYEMPNYRGRQYLLRPQEYRRYHDWGAVDAKAGSLRRVVDLY 174


RESULT 9
C24060
gamma-crystallin 2-1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 21-Nov-1997
C;Accession: C24060
R;den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A;Reference number: A92927; MUID:87060933; PMID:3783678
A;Accession: C24060
A;Molecule type: DNA
A;Residues: 1-174 <DEN>
A;Cross-references: GB:M19359; GB:M19354; GB:M19360; GB:M19361; GB:M19362; GB:M19363
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

  Query Match           70.0%;  Score 775.5;  DB 2;  Length 174;
  Best Local Similarity  77.7%;  Pred. No. 5.8e-66;
  Matches  136;  Conservative   19;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| | |:||||||| |||||||:|| ||||||||||||||||||||
Db          1 MGKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:|||||||||| |||: |||:||::| :| | |:::|| :||||||:
Db         61 GDYPDYQQWMGFSDSIRSCRLIP-HTGSHRMRLYEKEDHKGVMMELSEDCSCIQDRFHLS 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              || ||:|||| |||||||:|||||||||| ||||| ||||::|| ||||||:| |
Db        120 EVRSLHVLEGCWVLYEMPNYRGRQYLLRPQEYRRYHDWGAVDAKAGSLRRVVDLY 174


RESULT 10
A24060
gamma-crystallin 1-1 - rat
N;Alternate names: gamma-A-crystallin
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: A24060; I60312
R;den Dunnen, J.T.; Moormann, R.J.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
J. Mol. Biol. 189, 37-46, 1986
A;Title: Concerted and divergent evolution within the rat gamma-crystallin gene family.
A;Reference number: A92927; MUID:87060933; PMID:3783678
A;Accession: A24060
A;Molecule type: DNA
A;Residues: 1-174 <DEN>
A;Cross-references: UNIPROT:P10065; GB:M19359; NID:g203626; PIDN:AAA40981.1; PID:g203627,
R;Den Dunnen, J.T.; van Neck, J.W.; Cremers, F.P.M.; Lubsen, N.H.; Schoenmakers, J.G.G.
Gene 78, 201-213, 1989
A;Title: Nucleotide sequence of the rat gamma-crystallin gene region and comparison with
A;Reference number: I60312; MUID:89378747; PMID:2777080
A;Accession: I60312
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-174 <DE2>
A;Cross-references: GB:M19359; NID:g203626; PIDN:AAA40981.1; PID:g203627
C;Genetics:
A;Gene: CRY-gamma-A
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

  Query Match           69.9%;  Score 774.5;  DB 2;  Length 174;
  Best Local Similarity  77.7%;  Pred. No. 7.2e-66;
  Matches  136;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| | |:||||||| |||||||||| ||||||||||||||:||||||
Db          1 MGKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||||||| || :| : |:|:|||||:|| :||:|:||  : ||| |
Db         61 GDYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:|::|||||||||||:||||||||:|||| |||||:|||||||||| |
Db        120 EIYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 174


RESULT 11
CYRTG1
gamma-crystallin 1 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 13-Jun-1983 #sequence_revision 13-Jun-1983 #text_change 09-Jul-2004
C;Accession: A02930
R;Moormann, R.J.M.; den Dunnen, J.T.; Bloemendal, H.; Schoenmakers, J.G.G.
Proc. Natl. Acad. Sci. U.S.A. 79, 6876-6880, 1982
```

```
A;Title: Extensive intragenic sequence homology in two distinct rat lens gamma-crystalli
A;Reference number: A93934; MUID:83091061; PMID:6294661
A;Accession: A02930
A;Molecule type: mRNA
A;Residues: 1-173 <MOO>
A;Cross-references: UNIPROT:P02528; GB:J00716; NID:g203634; PIDN:AAA40987.1; PID:g203635
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;1-39/Domain: crystallin repeat <GK1>
F;40-82/Domain: crystallin repeat <GK2>
F;87-127/Domain: crystallin repeat <GK3>
F;128-167/Domain: crystallin repeat <GK4>

  Query Match             69.8%;  Score 773.5;  DB 1;  Length 173;
  Best Local Similarity   77.0%;  Pred. No. 8.9e-66;
  Matches  134;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |::|  |||||||||||:|| ||||||||:||: | |||||||
Db          1 GKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:||:|||||| |: : |:|||||:|:|||| ||||||| |||||| ::
Db         61 DYPDYQQWMGFSDSVRSCRLIP-HSSSHRIRIYEREDYRGQMVEITDDCPHLQDRFHFSD 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              || :|:|| |||||||:||||||||||||||| |||||||:||||||:||||
Db        120 FHSFHVMEGYWVLYEMPNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRIMDFY 173


RESULT 12
S04266
gamma-crystallin IV - bovine
C;Species: Bos primigenius taurus (cattle)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 17-Mar-1999
C;Accession: S04266; C60890
R;White, H.E.; Driessen, H.P.C.; Slingsby, C.; Moss, D.S.; Lindley, P.F.
J. Mol. Biol. 207, 217-235, 1989
A;Title: Packing interactions in the eye-lens. Structural analysis, internal symmetry an
A;Reference number: S04265; MUID:89293855; PMID:2738925
A;Accession: S04266
A;Molecule type: protein
A;Residues: 1-173 <WHI>
R;Chiou, S.H.; Azari, P.; Himmel, M.E.
J. Protein Chem. 7, 67-80, 1988
A;Title: Physicochemical characterization of gamma-crystallins from bovine lens-hydrodyn
A;Reference number: A60890; MUID:89351571; PMID:3255364
A;Accession: C60890
A;Molecule type: protein
A;Residues: 1-25 <CHI>
A;Note: 14-His, 15-Arg, 22-Cys, 23-Pro, and 24-Asp were also found and may indicate cont
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens

  Query Match             69.5%;  Score 770.5;  DB 2;  Length 173;
  Best Local Similarity   77.0%;  Pred. No. 1.7e-65;
  Matches  134;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |::|  |||||||||||||| |||||||:||: | |||||||
Db          1 GKITFYEDRGFQGRHYECSSDHSNLQPYFSRCNSIRVDSGCWMLYEQPNFTGCQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              ||||||||||:||:||||||| || : |:|||||:|:|||| |||:|| ||||||| ::
Db         61 DYPDYQQWMGFSDSVRSCRLIP-HTSSHRLRIYEREDYRGQMVEITEDCSSLQDRFHFSD 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :|| :|:|| |||||||:||||||||||:||||||| ||:|||||| :|||
Db        120 IHSFHVMEGYWVLYEMPNYRGRQYLLRPGDYRRYLDWGAANARVGSLRRAVDFY 173


RESULT 13
CYMSG1
gamma-crystallin 1 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C;Accession: A02932
R;Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Tsui
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A;Title: gamma-crystallin family of the mous lens: structural and evolutionary relationsh
A;Reference number: A94021; MUID:85088487; PMID:6096855
A;Accession: A02932
A;Molecule type: mRNA
A;Residues: 1-174 <BRE>
A;Cross-references: UNIPROT:P04342; GB:K02583; NID:g192778; PIDN:AAA37475.1; PID:g309197
C;Comment: There are at least seven different gamma crystallins identified in mouse lens
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;88-128/Domain: crystallin repeat <GK3>
F;129-168/Domain: crystallin repeat <GK4>

  Query Match             69.5%;  Score 770.5;  DB 1;  Length 174;
  Best Local Similarity   76.6%;  Pred. No. 1.7e-65;
  Matches  134;  Conservative   21;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  ||||||| |||:|| |||||||:||: | ||||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSHCNSVRVDSGCWMLYEQPNFAGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:||||||| | |: |:|:|||:::|||: | |:||||||||||
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HAGSHRIRLYEREEYRGQVIEFTEDCPSLQDRFHFN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |::||||||| |||||:| :|||||||||||||||| ||||||||||||||||||
Db        120 EIYSLNVLEGCWVLYDMTNYRGRQYLLRPGEYRRYHDWGAMNAKVGSLRRVMDFY 174


RESULT 14
JS0596
gamma-E-crystallin - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 21-Nov-1997
C;Accession: JS0596; S18835
R;Graw, J.; Coban, L.; Liebstein, A.; Werner, T.
Gene 104, 265-270, 1991
A;Title: Murine gamma E-crystallin is distinct from murine gamma 2-crystallin.
A;Reference number: JS0596; MUID:92009223; PMID:1916296
A;Accession: JS0596
A;Molecule type: DNA
A;Residues: 1-174 <GRA>
A;Cross-references: EMBL:X57855
A;Note: this sequence differs two residues from that of rat
C;Genetics:
A;Gene: gamma-E-cry
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication

  Query Match             69.3%;  Score 767.5;  DB 2;  Length 174;
  Best Local Similarity   76.0%;  Pred. No. 3.3e-65;
  Matches  133;  Conservative   22;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  |||||||||||:|| ||||||||:||: | ||||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:||||||| |: : |::||||:|:|||| |||||| |||||| :
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HSSSHRIKIYEREDYRGQMVEITDDCSHLQDRFHFS 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              : || :|:|| |||||||:||||||||||||||| |||||||:|||||:||||
Db        120 DFHSFHVMEGYWVLYEMPNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRIMDFY 174
```

RESULT 15
CYMSG4
gamma-crystallin 4 - mouse
C;Species: Mus musculus (house mouse)
C;Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C;Accession: A02935; I49613
R;Breitman, M.L.; Lok, S.; Wistow, G.; Piatigorsky, J.; Treton, J.A.; Gold, R.J.M.; Tsui
Proc. Natl. Acad. Sci. U.S.A. 81, 7762-7766, 1984
A;Title: gamma-crystallin family of the mous lens: structural and evolutionary relations
A;Reference number: A94021; MUID:85088487; PMID:6096855
A;Accession: A02935
A;Molecule type: mRNA
A;Residues: 1-174 <BRE>
A;Cross-references: UNIPROT:P04345
A;Note: the authors translated the codon ATC for residue 36 as Val
R;Lok, S.; Tsui, L.C.; Shinohara, T.; Piatigorsky, J.; Gold, R.; Breitman, M.
Nucleic Acids Res. 12, 4517-4529, 1984
A;Title: Analysis of the mouse gamma-crystallin gene family: assignment of multiple cDNA
A;Reference number: I48353; MUID:84247318; PMID:6330674
A;Accession: I49613
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-174 <RES>
A;Cross-references: GB:K02587; NID:g192771; PIDN:AAA37473.1; PID:g387135
C;Comment: There are at least seven different gamma crystallins identified in mouse lens
C;Genetics:
A;Introns: 3/3; 84/3
C;Superfamily: beta-crystallin
C;Keywords: duplication; eye lens
F;2-40/Domain: crystallin repeat <GK1>
F;41-83/Domain: crystallin repeat <GK2>
F;88-128/Domain: crystallin repeat <GK3>
F;129-168/Domain: crystallin repeat <GK4>

Query Match 69.1%; Score 765.5; DB 1; Length 174;
Best Local Similarity 77.1%; Pred. No. 5.1e-65;
Matches 135; Conservative 19; Mismatches 20; Indels 1; Gaps 1;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | |||||||  | |:|||||||| |||||||||| ||||||||||||||:||||||
Db          1 MGKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:|||||| || :| : |:|:|||||:|| :||: |||  : ||| |
Db         61 GDYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELMDDCSCIHDRFRLH 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |::|::|||| |||||||:||||||||||:|||| |||||:|||||||||| |
Db        120 EIYSMHVLEGCWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 174
```

Search completed: January 14, 2005, 12:07:11
Job time : 17 secs

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 12:03:35 ; Search time 194 Seconds
(without alignments)
584.272 Million cell updates/sec

Title: US-10-030-605A-21
Perfect score: 1108
Sequence: 1 MGRIKFKEDRGFQGHYYSCN.....PNSSSVDKLAAALEHHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_02:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	938	84.7	174	1	CRGB_BOVIN	P02526 bos taurus
2	876	79.1	174	1	CRGB_RAT	P10066 rattus norv
3	859	77.5	174	1	CRGB_MOUSE	P04344 mus musculu
4	856	77.3	175	2	Q6PHP7	Q6php7 mus musculu
5	856	77.3	175	2	AAH56455	Aah56455 mus muscu
6	825	74.5	174	1	CRGB_HUMAN	P07316 homo sapien
7	801	72.3	174	1	CRGA_BOVIN	P02527 bos taurus
8	778.5	70.3	173	1	CRGD_BOVIN	P08209 bos taurus
9	778.5	70.3	173	1	CRGD_RAT	P10067 rattus norv
10	776.5	70.1	174	2	Q6PGI0	Q6pgi0 mus musculu
11	776.5	70.1	174	2	AAH57013	Aah57013 mus muscu
12	773.5	69.8	173	1	CRGE_RAT	P02528 rattus norv
13	772.5	69.7	173	1	CRGF_BOVIN	P23005 bos taurus
14	771.5	69.6	173	1	CRGC_RAT	P02529 rattus norv
15	769.5	69.4	173	1	CRGA_RAT	P10065 rattus norv
16	765.5	69.1	173	1	CRGD_MOUSE	P04342 mus musculu
17	762.5	68.8	173	1	CRGE_MOUSE	Q03740 mus musculu
18	760.5	68.6	173	1	CRGA_MOUSE	P04345 mus musculu
19	757.5	68.4	173	1	CRGF_RAT	P10068 rattus norv
20	757.5	68.4	174	2	AAH56453	Aah56453 mus muscu
21	752.5	67.9	173	1	CRGF_MOUSE	Q9cxv3 mus musculu
22	751.5	67.8	173	1	CRGE_BOVIN	Q28088 bos taurus
23	747.5	67.5	173	1	CRGC_MOUSE	Q61597 mus musculu
24	722.5	65.2	173	1	CRGD_HUMAN	P07320 homo sapien
25	721.5	65.1	173	1	CRGC_HUMAN	P07315 homo sapien
26	710.5	64.1	173	1	CRGA_HUMAN	P11844 homo sapien
27	681	61.5	175	1	CRG2_CHICO	P48649 chiloscylli
28	671	60.6	175	2	Q6DKC9	Q6dkc9 xenopus lae
29	654	59.0	175	2	Q7SZA6	Q7sza6 xenopus lae
30	651	58.8	172	1	CRBS_CHICO	P48646 chiloscylli
31	649	58.6	175	2	Q6DJC9	Q6djc9 xenopus lae
32	647	58.4	175	2	Q6DER7	Q6der7 xenopus tro
33	645	58.2	175	1	CRG3_XENLA	P55940 xenopus lae
34	644	58.1	174	2	Q9PSY5	Q9psy5 xenopus lae
35	641	57.9	175	2	Q8QFU3	Q8qfu3 cynops pyrr
36	639	57.7	169	1	CRG2_RANTE	P02531 rana tempor
37	635	57.3	177	2	O93615	O93615 xenopus lae
38	631	56.9	175	1	CRG2_XENLA	Q91724 xenopus lae
39	627	56.6	177	1	CRG1_RANCA	Q91320 rana catesb
40	621	56.0	175	1	CRG1_XENLA	Q06254 xenopus lae
41	616	55.6	177	1	CRG2_RANCA	Q91321 rana catesb
42	610	55.1	172	1	CRG4_XENLA	P55941 xenopus lae
43	605	54.6	172	1	CRBT_CHICO	P48647 chiloscylli
44	561	50.6	133	1	CRG1_RANTE	P02530 rana tempor
45	549.5	49.6	173	1	CRBS_CYPCA	P10112 cyprinus ca

ALIGNMENTS

RESULT 1
CRGB_BOVIN
ID CRGB_BOVIN STANDARD; PRT; 174 AA.
AC P02526;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 01-OCT-2004 (Rel. 45, Last annotation update)
DE Gamma crystallin B (Gamma crystallin II).
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87270760; PubMed=3606621;
RA Hay R.E., Woods W.D., Church R.L., Petrash J.M.;
RT "cDNA clones encoding bovine gamma-crystallins.";
RL Biochem. Biophys. Res. Commun. 146:332-338(1987).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lens;
RX MEDLINE=85026663; PubMed=6092016;
RA Bhat S.P., Spector A.;
RT "Complete nucleotide sequence of a cDNA derived from calf lens gamma-
RT crystallin mRNA: presence of Alu I-like DNA sequences.";
RL DNA 3:287-295(1984).
RN [3]
RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS), AND REVISIONS.
RX MEDLINE=81123111; PubMed=7464942;
RA Blundell T.L., Lindley P., Miller L., Moss D., Slingsby C., Tickle I.,
RA Turnell B., Wistow G.;
RT "The molecular structure and stability of the eye lens: X-ray analysis
RT of gamma-crystallin II.";
RL Nature 289:771-777(1981).
RN [4]
RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS).
RX MEDLINE=84036195; PubMed=6631960;
RA Wistow G., Turnell B., Summers L., Slingsby C., Moss D., Miller L.,
RA Lindley P., Blundell T.L.;
RT "X-ray analysis of the eye lens protein gamma-II crystallin at 1.9-A
RT resolution.";
RL J. Mol. Biol. 170:175-202(1983).
RN [5]
RP X-RAY CRYSTALLOGRAPHY (1.47 ANGSTROMS).
RA Najmudin S., Nalini V., Dreissen H.P.C., Slingsby C., Blundell T.L.,
RA Moss D.S., Lindley P.F.;
RT "Structure of the bovine eye lens protein gammaB (gammaII)-crystallin
RT at 1.47 A.";
RL Acta Crystallogr. D 49:223-233(1993).
RN [6]
RP X-RAY CRYSTALLOGRAPHY (1.2 ANGSTROMS).
RA Kumaraswamy V.S., Lindley P.F., Slingsby C., Glover I.D.;
RT "An eye lens protein-water structure: 1.2-A resolution structure of

```
RT   gammaB-crystallin at 150 K.";
RL   Acta Crystallogr. D 52:611-622(1996).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY.
RX   MEDLINE=98200466; PubMed=9541393;
RA   Palme S., Jaenicke R., Slingsby C.;
RT   "X-ray structures of three interface mutants of gammaB-crystallin from
RT   bovine eye lens.";
RL   Protein Sci. 7:611-618(1998).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF MUTANT ALA-56.
RX   MEDLINE=98308013; PubMed=9642083;
RA   Palme S., Jaenicke R., Slingsby C.;
RT   "Unusual domain pairing in a mutant of bovine lens gammaB-
RT   crystallin.";
RL   J. Mol. Biol. 279:1053-1059(1998).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X01036; CAA25518.1; -.
DR   EMBL; M16894; AAA30476.1; -.
DR   PIR; A29655; CYBOG.
DR   PDB; 1AMM; X-ray; @=1-174.
DR   PDB; 1DSL; X-ray; @=87-174.
DR   PDB; 1GAM; X-ray; A/B=87-172.
DR   PDB; 1GCS; X-ray; @=1-174.
DR   PDB; 1I5I; X-ray; A=1-174.
DR   PDB; 4GCR; X-ray; @=1-174.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   3D-structure; Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     87       Connecting peptide.
FT   DOMAIN       88    128       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      129    171       Beta/gamma crystallin 'Greek key' 4.
FT   CONFLICT    119    119       T -> S (in Ref. 2).
FT   STRAND        2      8
FT   HELIX         9     11
FT   STRAND       12     18
FT   STRAND       22     22
FT   TURN         26     28
FT   STRAND       34     39
FT   STRAND       42     47
FT   TURN         48     50
FT   STRAND       51     57
FT   STRAND       60     62
FT   HELIX        65     68
FT   TURN         69     69
FT   STRAND       77     80
FT   STRAND       89     95
FT   TURN         96     98
FT   STRAND       99    105
FT   STRAND      109    109
FT   HELIX       112    116
FT   STRAND      120    120
FT   STRAND      123    128
FT   STRAND      131    136
FT   TURN        137    139
FT   STRAND      140    146
FT   STRAND      149    151
FT   HELIX       154    157
FT   TURN        158    158
FT   STRAND      163    163
FT   STRAND      166    169
SQ   SEQUENCE   174 AA;  20965 MW;  8E404878CA2150A3 CRC64;

  Query Match             84.7%;  Score 938;  DB 1;  Length 174;
  Best Local Similarity   96.0%;  Pred. No. 9.1e-79;
  Matches  167;  Conservative    2;  Mismatches    5;  Indels    0;  Gaps    0;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| | |:||||||||||||||||| ||||||||||||||||||||
Db          1 GKITFYEDRGFQGHCYECSSDCPNLQPYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 120

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 174


RESULT 2
CRGB_RAT
ID   CRGB_RAT       STANDARD;      PRT;   174 AA.
AC   P10066;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin B (Gamma crystallin 1-2).
GN   Name=Crygb;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236314; PubMed=1970548;
RA   den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA   Schoenmakers J.G.G.;
RT   "Isolation and characterization of the rat glutamine synthetase-
RT   encoding gene.";
RL   Gene 87:225-232(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87060933; PubMed=3783678;
RA   den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT   "Concerted and divergent evolution within the rat gamma-crystallin
RT   gene family.";
RL   J. Mol. Biol. 189:37-46(1986).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- MISCELLANEOUS: There are six different gamma crystallins
CC       identified in rat lens.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
```

```
DR   EMBL; M19359; AAA40982.1; -.
DR   PIR; B24060; B24060.
DR   HSSP; P02526; 1AMM.
DR   RGD; 2420; Crygb.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     87       Connecting peptide.
FT   DOMAIN       88    128       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      129    171       Beta/gamma crystallin 'Greek key' 4.
FT   CONFLICT     15     15       C -> S (in Ref. 2).
SQ   SEQUENCE   174 AA;  20957 MW;  076955AA5A94C70C CRC64;

  Query Match             79.1%;  Score 876;  DB 1;  Length 174;
  Best Local Similarity   88.5%;  Pred. No. 5e-73;
  Matches  154;  Conservative    9;  Mismatches   11;  Indels    0;  Gaps    0;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||:|| ||||||||||||||||||||
Db          1 GKITFFEDRGFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:||||||||||:||:||||||||||||||||||| ||||||||:|
Db         61 DYPDYQQWMGFSDSIRSCRLIPQHSGTYRMRIYERDDFRGQMSEITDDCLSLQDRFHLSE 120

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :|||||:|| |||||||||||||||||||||||||||||| |||||| |||||||
Db        121 IHSLNVMEGCWVLYEMPSYRGRQYLLRPGEYRRYLDWGAANAKVGSFRRVMDFY 174


RESULT 3
CRGB_MOUSE
ID   CRGB_MOUSE     STANDARD;      PRT;   174 AA.
AC   P04344; Q61593;
DT   20-MAR-1987 (Rel. 04, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Gamma crystallin B (Gamma crystallin 3).
GN   Name=Crygb;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=102 X C3H; TISSUE=Liver;
RX   MEDLINE=94123992; PubMed=8293998;
RA   Graw J., Liebstein A., Pietrowski D., Schmitt-John T., Werner T.;
RT   "Genomic sequences of murine gamma B- and gamma C-crystallin-encoding
RT   genes: promoter analysis and complete evolutionary pattern of mouse,
RT   rat and human gamma-crystallins.";
RL   Gene 136:145-156(1993).
RN   [2]
RP   SEQUENCE OF 37-174 FROM N.A.
RX   MEDLINE=85088487; PubMed=6096855;
RA   Breitman M.L., Lok S., Wistow G., Piatigorsky J., Treton J.A.,
RA   Gold R.J.M., Tsui L.-C.;
RT   "Gamma-crystallin family of the mouse lens: structural and
RT   evolutionary relationships.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:7762-7766(1984).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- MISCELLANEOUS: There are six different gamma crystallins
CC       identified in mouse lens.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; K02585; AAA37474.1; -.
DR   EMBL; Z22573; CAA80296.1; -.
DR   PIR; I48359; CYMSG3.
DR   HSSP; P02526; 1AMM.
DR   MGD; MGI:88522; Crygb.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0       By similarity.
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     87       Connecting peptide.
FT   DOMAIN       88    128       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      129    171       Beta/gamma crystallin 'Greek key' 4.
SQ   SEQUENCE   174 AA;  21007 MW;  9D206049E8F7327D CRC64;

  Query Match             77.5%;  Score 859;  DB 1;  Length 174;
  Best Local Similarity   86.2%;  Pred. No. 1.9e-71;
  Matches  150;  Conservative   11;  Mismatches   13;  Indels    0;  Gaps    0;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||| |||  | |:|||||| |||||||:|| ||||||||||||||||||||
Db          1 GKITFFEDRSFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              :|||||||||:|||||||||||:||:|||||||:||||||||||| ||||||| :|
Db         61 EYPDYQQWMGFSDSIRSCRLIPQHSGTYRMRIYEKDDFRGQMSEITDDCLSLQDRFHFSE 120

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :|||||:|| |||||||||||||||||||||||||||||| |||||| |||||||
Db        121 IHSLNVMEGCWVLYEMPSYRGRQYLLRPGEYRRYLDWGAANAKVGSFRRVMDFY 174


RESULT 4
Q6PHP7
ID   Q6PHP7      PRELIMINARY;      PRT;   175 AA.
AC   Q6PHP7;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Crystallin, gamma B.
GN   Name=Crygb;
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
```

```
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens (By similarity).
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs (By similarity).
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
DR   EMBL; BC056455; AAH56455.1; -.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
SQ   SEQUENCE   175 AA;  21086 MW;  1E39B15A0AAB175D CRC64;

  Query Match             77.3%;  Score 856;  DB 2;  Length 175;
  Best Local Similarity   85.7%;  Pred. No. 3.6e-71;
  Matches  150;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||| ||| | |:||||||| ||||||:|| ||||||||||||||||||||
Db          1 MGKITFFEDRSFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |:||||||||||:|||||| |||||:||:||||||:||||||||||||||| |||||||
Db         61 GEYPDYQQWMGFSDSIRSCCLIPQHSGTYRMRIYEKDDFRGQMSEITDDCLSLQDRFHFS 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:|||||:|| |||||||||||||||||||||||||||||| |||||| |||||||
Db        121 EIHSLNVMEGCWVLYEMPSYRGRQYLLRPGEYRRYLDWGAANAKVGSFRRVMDFY 175


RESULT 5
AAH56455
ID   AAH56455    PRELIMINARY;      PRT;   175 AA.
AC   AAH56455;
DT   10-MAY-2004 (TrEMBLrel. 27, Created)
DT   10-MAY-2004 (TrEMBLrel. 27, Last sequence update)
DT   10-MAY-2004 (TrEMBLrel. 27, Last annotation update)
DE   Crystallin, gamma B.
GN   CRYGB.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC056455; AAH56455.1; -.
SQ   SEQUENCE   175 AA;  21086 MW;  1E39B15A0AAB175D CRC64;

  Query Match             77.3%;  Score 856;  DB 2;  Length 175;
  Best Local Similarity   85.7%;  Pred. No. 3.6e-71;
  Matches  150;  Conservative   11;  Mismatches   14;  Indels    0;  Gaps    0;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||| ||| | |:||||||| ||||||:|| ||||||||||||||||||||
Db          1 MGKITFFEDRSFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |:||||||||||:|||||| |||||:||:||||||:||||||||||||||| |||||||
Db         61 GEYPDYQQWMGFSDSIRSCCLIPQHSGTYRMRIYEKDDFRGQMSEITDDCLSLQDRFHFS 120

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:|||||:|| |||||||||||||||||||||||||||||| |||||| |||||||
Db        121 EIHSLNVMEGCWVLYEMPSYRGRQYLLRPGEYRRYLDWGAANAKVGSFRRVMDFY 175


RESULT 6
CRGB_HUMAN
ID   CRGB_HUMAN     STANDARD;      PRT;   174 AA.
AC   P07316;
DT   01-APR-1988 (Rel. 07, Created)
DT   01-APR-1988 (Rel. 07, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin B (Gamma crystallin 1-2).
GN   Name=CRYGB; Synonyms=CRYG2;
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=86056977; PubMed=4065573;
RA   den Dunnen J.T., Moormann R.J.M., Cremers F.P.M., Schoenmakers J.G.G.;
RT   "Two human gamma-crystallin genes are linked and riddled with Alu-
RT   repeats.";
RL   Gene 38:197-204(1985).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89378747; PubMed=2777080;
RA   den Dunnen J.T., van Neck J.W., Cremers F.P., Lubsen N.H.,
RA   Schoenmakers J.G.;
RT   "Nucleotide sequence of the rat gamma-crystallin gene region and
RT   comparison with an orthologous human region.";
RL   Gene 78:201-213(1989).
RN   [3]
RP   3D-STRUCTURE MODELING.
RX   MEDLINE=22395646; PubMed=12507494; DOI=10.1016/S0006-291X(02)02895-4;
RA   Salim A., Zaidi Z.H.;
RT   "Homology models of human gamma-crystallins: structural study of the
RT   extensive charge network in gamma-crystallins.";
RL   Biochem. Biophys. Res. Commun. 300:624-630(2003).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
```

```
CC       the vertebrate eye lens.
CC   -!- SUBUNIT: Monomer (By similarity).
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11971; AAA52113.1; -.
DR   EMBL; M11970; AAA52113.1; JOINED.
DR   EMBL; M19364; AAA52109.1; -.
DR   PIR; A24520; CYHUG1.
DR   PDB; 1LEU; Model; A=1-174.
DR   Genew; HGNC:2409; CRYGB.
DR   MIM; 123670; -.
DR   GO; GO:0005212; F:structural constituent of eye lens; NAS.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   3D-structure; Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     87       Connecting peptide.
FT   DOMAIN       88    128       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      129    171       Beta/gamma crystallin 'Greek key' 4.
SQ   SEQUENCE   174 AA;  20776 MW;  6573BE9001EBBFCC CRC64;

  Query Match           74.5%;  Score 825;  DB 1;  Length 174;
  Best Local Similarity   81.6%;  Pred. No. 2.6e-68;
  Matches  142;  Conservative   16;  Mismatches   16;  Indels    0;  Gaps    0;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||| |||   | | :||||||||||||||||| |||||:||||||||||||||
Db          1 GKITFYEDRAFQGRSYECTTDCPNLQPYFSRCNSIRVESGCWMIYERPNYQGHQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              :|||||||| :|||||| ||| |:| :||:||:||: |||||||:|||| |:||||||||
Db         61 EYPDYQQWMGLSDSIRSCCLIPPHSGAYRMKIYDRDELRGQMSELTDDCLSVQDRFHLTE 120

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :|||||||||:||||||||||||||||||:||||| |||||||||||| |
Db        121 IHSLNVLEGSWILYEMPNYRGRQYLLRPGEYRRFLDWGAPNAKVGSLRRVMDLY 174


RESULT 7
CRGA_BOVIN
ID   CRGA_BOVIN     STANDARD;      PRT;   174 AA.
AC   P02527;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   05-JUL-2004 (Rel. 44, Last annotation update)
DE   Gamma crystallin A (Gamma crystallin IVB).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE.
RX   MEDLINE=73054483; PubMed=4674126;
RA   Croft L.R.;
RT   "The amino acid sequence of gamma-crystallin (fraction II) from calf
RT   lens.";
RL   Biochem. J. 128:961-970(1972).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
DR   HSSP; P02526; 1DSL.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   Direct protein sequencing; Eye lens protein; Multigene family; Repeat.
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     87       Connecting peptide.
FT   DOMAIN       88    128       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      129    171       Beta/gamma crystallin 'Greek key' 4.
SQ   SEQUENCE   174 AA;  21003 MW;  E39093A105E2CEEF CRC64;

  Query Match           72.3%;  Score 801;  DB 1;  Length 174;
  Best Local Similarity   82.8%;  Pred. No. 4.3e-66;
  Matches  144;  Conservative   13;  Mismatches   17;  Indels    0;  Gaps    0;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||| | |:|:    ||||||||||||     | :|:||:|:|||| |:||
Db          1 GKITFYEDRGFQGHCYQCSSNNCLQQPYFSRCNSIRVDVHSWFVYQRPDYRGHQYMLQRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              :|| | |||||:|||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 NYPQYGQWMGFDDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 120

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |:|: |||||||:|||||||||||||||||||||||||||||||||||||||||
Db        121 VNSVRVLEGSWVIYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 174


RESULT 8
CRGD_BOVIN
ID   CRGD_BOVIN     STANDARD;      PRT;   173 AA.
AC   P08209; Q28089;
DT   01-AUG-1988 (Rel. 08, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin D (Gamma crystallin IIIB).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC   TISSUE=Lens;
RX   MEDLINE=95010404; PubMed=7925695;
RA   Hay R.E., Andley U.P., Petrash J.M.;
RT   "Expression of recombinant bovine gamma B-, gamma C- and gamma D-
RT   crystallins and correlation with native proteins.";
RL   Exp. Eye Res. 58:573-584(1994).
RN   [2]
RP   SEQUENCE OF 1-156 FROM N.A.
RC   TISSUE=Lens;
RX   MEDLINE=87270760; PubMed=3606621;
RA   Hay R.E., Woods W.D., Church R.L., Petrash J.M.;
RT   "cDNA clones encoding bovine gamma-crystallins.";
RL   Biochem. Biophys. Res. Commun. 146:332-338(1987).
RN   [3]
RP   X-RAY CRYSTALLOGRAPHY (1.95 ANGSTROMS).
RA   Chirgadze Y.N., Driessen H.P.C., Wright G., Slingsby C., Hay R.E.,
RA   Lindley P.F.;
RT   "Structure of bovine eye lens gammaD (gammaIIIb)-crystallin at 1.95
```

```
RT   A.";
RL   Acta Crystallogr. D 52:712-721(1996).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; L27070; AAA98995.1; -.
DR   EMBL; M16895; AAB59282.1; -.
DR   PDB; 1ELP; X-ray; A/B=1-173.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   3D-structure; Direct protein sequencing; Eye lens protein;
KW   Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     86       Connecting peptide.
FT   DOMAIN       87    127       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      128    170       Beta/gamma crystallin 'Greek key' 4.
FT   CONFLICT     75     75       V -> I (in Ref. 2).
FT   STRAND        2      8
FT   HELIX         9     11
FT   STRAND       12     18
FT   STRAND       22     22
FT   TURN         26     28
FT   STRAND       34     39
FT   STRAND       42     45
FT   TURN         49     50
FT   STRAND       54     57
FT   STRAND       60     62
FT   HELIX        65     68
FT   TURN         69     69
FT   STRAND       77     80
FT   STRAND       88     94
FT   TURN         95     97
FT   STRAND       98    104
FT   STRAND      108    108
FT   HELIX       111    114
FT   TURN        115    115
FT   STRAND      119    119
FT   STRAND      122    127
FT   STRAND      130    135
FT   TURN        136    138
FT   STRAND      139    145
FT   STRAND      148    150
FT   HELIX       153    156
FT   TURN        157    157
FT   STRAND      162    162
FT   STRAND      165    168
SQ   SEQUENCE   173 AA;  20735 MW;  D9D853EB3B3F7B5F CRC64;

  Query Match             70.3%;  Score 778.5;  DB 1;  Length 173;
  Best Local Similarity   78.2%;  Pred. No. 5.2e-64;
  Matches  136;  Conservative   17;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |:||  ||||| ||||:|| |||||:||:||| | |||||||
Db          1 GKITFYEDRGFQGRHYECSSDHSNLQPYLGRCNSVRVDSGCWMIYEQPNYLGPQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||| |||:||||||| | |: |:|:|||:|:|||| |||:|| ||||||| |
Db         61 DYPDYQQWMGLNDSVRSCRLIP-HAGSHRLRLYEREDYRGQMIEITEDCSSLQDRFHFNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :|||||||||||||:|:|||||||||||||||| ||||||||||||||:| |
Db        120 IHSLNVLEGSWVLYELPNYRGRQYLLRPGEYRRYHDWGAMNAKVGSLRRVIDIY 173


RESULT 9
CRGD_RAT
ID   CRGD_RAT       STANDARD;      PRT;   173 AA.
AC   P10067;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-MAR-1989 (Rel. 10, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin D (Gamma crystallin 2-2).
GN   Name=Crygd;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236314; PubMed=1970548;
RA   den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA   Schoenmakers J.G.G.;
RT   "Isolation and characterization of the rat glutamine synthetase-
RT   encoding gene.";
RL   Gene 87:225-232(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87060933; PubMed=3783678;
RA   den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT   "Concerted and divergent evolution within the rat gamma-crystallin
RT   gene family.";
RL   J. Mol. Biol. 189:37-46(1986).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Sprague-Dawley; TISSUE=Lens;
RX   MEDLINE=95151850; PubMed=7849105;
RA   Ooki K., Amuro N., Shimizu Y., Okazaki T.;
RT   "High level expression of rat gamma-D-crystallin in Escherichia
RT   coli.";
RL   Biochimie 76:398-403(1994).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- MISCELLANEOUS: There are six different gamma crystallins
CC       identified in rat lens.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M19359; AAA40984.1; -.
DR   EMBL; X57169; CAA40458.1; -.
DR   PIR; D24060; D24060.
DR   HSSP; P02528; 1A5D.
DR   RGD; 2422; Crygd.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
```

```
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     86       Connecting peptide.
FT   DOMAIN       87    127       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      128    170       Beta/gamma crystallin 'Greek key' 4.
SQ   SEQUENCE   173 AA;  20972 MW;  930B2D42D80F4117 CRC64;

  Query Match             70.3%;  Score 778.5;  DB 1;  Length 173;
  Best Local Similarity   78.2%;  Pred. No. 5.2e-64;
  Matches  136;  Conservative   19;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |::|  ||||||||||||:|| ||||||||:||: | |||||||
Db          1 GKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:||:|||||||  | |: |:|:|||:|:|||| | |:||||||||||  |
Db         61 DYPDYQQWMGFSDSVRSCRLIP-HAGSHRIRLYEREDYRGQMVEFTEDCPSLQDRFHFNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::||||||| |||||| :||||||||||||||| |||||||:||||||||||
Db        120 IYSLNVLEGCWVLYEMTNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRVMDFY 173


RESULT 10
Q6PGI0
ID   Q6PGI0      PRELIMINARY;      PRT;   174 AA.
AC   Q6PGI0;
DT   05-JUL-2004 (TrEMBLrel. 27, Created)
DT   05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT   05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE   Crystallin, gamma D.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens (By similarity).
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs (By similarity).
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
```

```
DR   EMBL; BC057013; AAH57013.1; -.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
SQ   SEQUENCE   174 AA;  21118 MW;  1AFC29C0269B35A1 CRC64;

  Query Match             70.1%;  Score 776.5;  DB 2;  Length 174;
  Best Local Similarity   77.1%;  Pred. No. 8e-64;
  Matches  135;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  ||||||||||||:|| ||||||||:||: | |||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:|||||||  | |: |:|:|||:::|||| | |:||||||||||
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HAGSHRIRLYEREEYRGQMIEFTEDCPSLQDRFHFN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |::||||||| |||:| :||||||||||||||| |||||||:||||||||||
Db        120 EIYSLNVLEGCWVLYDMTNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRVMDFY 174


RESULT 11
AAH57013
ID   AAH57013    PRELIMINARY;      PRT;   174 AA.
AC   AAH57013;
DT   02-MAR-2004 (TrEMBLrel. 27, Created)
DT   02-MAR-2004 (TrEMBLrel. 27, Last sequence update)
DT   02-MAR-2004 (TrEMBLrel. 27, Last annotation update)
DE   Crystallin, gamma D.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA   Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
RA   Jones S.J., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length human
RT   and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6; TISSUE=Brain;
RA   Strausberg R.;
RL   Submitted (AUG-2003) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC057013; AAH57013.1; -.
SQ   SEQUENCE   174 AA;  21118 MW;  1AFC29C0269B35A1 CRC64;

  Query Match             70.1%;  Score 776.5;  DB 2;  Length 174;
  Best Local Similarity   77.1%;  Pred. No. 8e-64;
  Matches  135;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;
```

```
Qy          1 MGRIKPKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| :| |::|  |||||||||||:|| ||||||||:||: | ||||||
Db          1 MGKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |||||||||||:||:|||||||  | |: |:|:|||:::|||| | |:||||||||||
Db         61 GDYPDYQQWMGFSDSVRSCRLIP-HAGSHRIRLYEREEYRGQMIEFTEDCPSLQDRFHFN 119

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              |::|||||| ||||:| :|||||||||||||| |||||||:||||||||||
Db        120 EIYSLNVLEGCWVLYDMTNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRVMDFY 174


RESULT 12
CRGE_RAT
ID   CRGE_RAT       STANDARD;      PRT;   173 AA.
AC   P02528;
DT   21-JUL-1986 (Rel. 01, Created)
DT   21-JUL-1986 (Rel. 01, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin E (Gamma crystallin 3-1) (Gamma-2).
GN   Name=Cryge;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236314; PubMed=1970548;
RA   den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA   Schoenmakers J.G.G.;
RT   "Isolation and characterization of the rat glutamine synthetase-
RT   encoding gene.";
RL   Gene 87:225-232(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=83091061; PubMed=6294661;
RA   Moormann R.J.M., den Dunnen J.T., Bloemendal H., Schoenmakers J.G.G.;
RT   "Extensive intragenic sequence homology in two distinct rat lens
RT   gamma-crystallin cDNAs suggests duplications of a primordial gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:6876-6880(1982).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=84114867; PubMed=6319707;
RA   Moormann R.J.M., den Dunnen J.T., Mulleners L., Andreoli P.,
RA   Bloemendal H., Schoenmakers J.G.G.;
RT   "Strict co-linearity of genetic and protein folding domains in an
RT   intragenically duplicated rat lens gamma-crystallin gene.";
RL   J. Mol. Biol. 171:353-368(1983).
RN   [4]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RC   TISSUE=Lens;
RX   MEDLINE=98035705; PubMed=9367641;
RA   Norledge B.V., Hay R.E., Bateman O.A., Slingsby C., Driessen H.P.C.;
RT   "Towards a molecular understanding of phase separation in the lens: a
RT   comparison of the X-ray structures of two high Tc gamma-crystallins,
RT   gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and
RT   gammaD.";
RL   Exp. Eye Res. 65:609-630(1997).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- MISCELLANEOUS: There are six different gamma crystallins
CC       identified in rat lens.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M19359; AAA40985.1; -.
DR   EMBL; J00716; AAA40987.1; -.
DR   EMBL; X00271; CAA25073.1; -.
DR   PIR; A02930; CYRTG1.
DR   PIR; I56381; I56381.
DR   PDB; 1A5D; X-ray; A/B=1-173.
DR   RGD; 2423; Cryge.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   3D-structure; Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     86       Connecting peptide.
FT   DOMAIN       87    127       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      128    170       Beta/gamma crystallin 'Greek key' 4.
FT   STRAND        2      8
FT   TURN          9     11
FT   STRAND       12     18
FT   STRAND       22     22
FT   HELIX        26     28
FT   STRAND       34     39
FT   STRAND       41     47
FT   TURN         48     50
FT   STRAND       51     57
FT   STRAND       60     62
FT   HELIX        65     68
FT   TURN         69     69
FT   STRAND       77     81
FT   STRAND       88     92
FT   HELIX        95     97
FT   STRAND      101    104
FT   STRAND      108    108
FT   HELIX       111    114
FT   STRAND      122    127
FT   STRAND      130    135
FT   TURN        136    138
FT   STRAND      139    145
FT   STRAND      148    150
FT   HELIX       153    156
FT   TURN        157    157
FT   STRAND      165    168
SQ   SEQUENCE   173 AA;  21132 MW;  3F3200E85CB61B02 CRC64;

  Query Match             69.8%;  Score 773.5;  DB 1;  Length 173;
  Best Local Similarity   77.0%;  Pred. No. 1.5e-63;
  Matches  134;  Conservative   21;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |::|  |||||||||||:|| ||||||||:||: | |||||||
Db          1 GKITFYEDRGFQGRHYECSTDHSNLQPYFSRCNSVRVDSGCWMLYEQPNFTGCQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:||:|||||||  |: : |:|||||:|:|||| ||||||| |||||| ::
Db         61 DYPDYQQWMGFSDSVRSCRLIP-HSSSHRIRIYEREDYRGQMVEITDDCPHLQDRFHFSD 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
               || :|:|| |||||||:|||||||||||||||| |||||||:|||||:||||
Db        120 FHSFHVMEGYWVLYEMPNYRGRQYLLRPGEYRRYHDWGAMNARVGSLRRIMDFY 173


RESULT 13
CRGF_BOVIN
ID   CRGF_BOVIN     STANDARD;      PRT;   173 AA.
AC   P23005;
DT   01-AUG-1991 (Rel. 19, Created)
```

```
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin F (Gamma crystallin IVA).
OS   Bos taurus (Bovine).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC   Bovinae; Bos.
OX   NCBI_TaxID=9913;
RN   [1]
RP   X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=89293855; PubMed=2738925;
RA   White H.E., Driessen H.P.C., Slingsby C., Moss D.S., Lindley P.F.;
RT   "Packing interactions in the eye-lens. Structural analysis, internal
RT   symmetry and lattice interactions of bovine gamma IVa-crystallin.";
RL   J. Mol. Biol. 207:217-235(1989).
RN   [2]
RP   SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
RX   MEDLINE=98035705; PubMed=9367641;
RA   Norledge B.V., Hay R.E., Bateman O.A., Slingsby C., Driessen H.P.C.;
RT   "Towards a molecular understanding of phase separation in the lens: a
RT   comparison of the X-ray structures of two high Tc gamma-crystallins,
RT   gammaE and gammaF, with two low Tc gamma-crystallins, gammaB and
RT   gammaD.";
RL   Exp. Eye Res. 65:609-630(1997).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
DR   PDB; 1A45; X-ray; @=1-173.
DR   PDB; 1M8U; X-ray; A=1-173.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   3D-structure; Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     86       Connecting peptide.
FT   DOMAIN       87    127       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      128    170       Beta/gamma crystallin 'Greek key' 4.
FT   STRAND        2      8
FT   TURN          9     11
FT   STRAND       12     18
FT   STRAND       22     22
FT   TURN         26     28
FT   STRAND       34     39
FT   STRAND       41     45
FT   HELIX        48     50
FT   STRAND       54     57
FT   STRAND       60     62
FT   TURN         65     69
FT   STRAND       77     81
FT   STRAND       88     92
FT   HELIX        95     97
FT   STRAND      101    104
FT   STRAND      108    108
FT   TURN         112    114
FT   STRAND      122    127
FT   STRAND      130    133
FT   STRAND      142    145
FT   STRAND      148    150
FT   HELIX       153    156
FT   TURN         157    157
FT   STRAND      165    168
SQ   SEQUENCE   173 AA;  20955 MW;  AC19C46CC323EC90 CRC64;

  Query Match             69.7%;  Score 772.5;  DB 1;  Length 173;
  Best Local Similarity   78.7%;  Pred. No. 1.9e-63;
  Matches  137;  Conservative   16;  Mismatches   20;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | ||||||| :| |:|| ||||||||||||||| ||||||||:||:|| |||||||
Db          1 GKITFYEDRGFQGRHYECSSDHSNLQPYFSRCNSIRVDSGCWMLYEQPNFQGPQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||| |||||||||| ||| |:|||||:|:|||| |||:|| || |||| :|
Db         61 DYPDYQQWMGLNDSIRSCRLIP-HTGSHRLRIYEREDYRGQMVEITEDCSSLHDRFHFSE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              :|| ||||| |||||| :|||||||||||:|||| |||| ||:|||||| :|||
Db        120 IHSFNVLEGWWVLYEMTNYRGRQYLLRPGDYRRYHDWGATNARVGSLRRAVDFY 173


RESULT 14
CRGC_RAT
ID   CRGC_RAT       STANDARD;      PRT;   173 AA.
AC   P02529;
DT   21-JUL-1986 (Rel. 01, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin C (Gamma crystallin 2-1).
GN   Name=Crygc;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236314; PubMed=1970548;
RA   den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA   Schoenmakers J.G.G.;
RT   "Isolation and characterization of the rat glutamine synthetase-
RT   encoding gene.";
RL   Gene 87:225-232(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87060933; PubMed=3783678;
RA   den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT   "Concerted and divergent evolution within the rat gamma-crystallin
RT   gene family.";
RL   J. Mol. Biol. 189:37-46(1986).
RN   [3]
RP   SEQUENCE OF 4-173 FROM N.A.
RX   MEDLINE=83091061; PubMed=6294661;
RA   Moormann R.J.M., den Dunnen J.T., Bloemendal H., Schoenmakers J.G.G.;
RT   "Extensive intragenic sequence homology in two distinct rat lens
RT   gamma-crystallin cDNAs suggests duplications of a primordial gene.";
RL   Proc. Natl. Acad. Sci. U.S.A. 79:6876-6880(1982).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- MISCELLANEOUS: There are six different gamma crystallins
CC       identified in rat lens.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ----------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ----------------------------------------------------------------------
DR   EMBL; M19359; AAA40983.1; -.
DR   EMBL; J00717; AAA40986.1; -.
DR   PIR; A02934; CYRTG2.
DR   PIR; I83432; I83432.
DR   HSSP; P02526; 1AMM.
DR   RGD; 2421; Crygc.
```

```
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     86       Connecting peptide.
FT   DOMAIN       87    127       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      128    170       Beta/gamma crystallin 'Greek key' 4.
FT   CONFLICT     15     15       C -> S (in Ref. 2 and 3).
FT   CONFLICT     35     35       I -> V (in Ref. 2 and 3).
FT   CONFLICT     83     83       H -> R (in Ref. 3).
FT   CONFLICT     87     87       H -> Q (in Ref. 3).
FT   CONFLICT    168    169       RV -> SA (in Ref. 3).
SQ   SEQUENCE   173 AA;  20819 MW;  1F7E7324A8EEF1DB CRC64;

  Query Match             69.6%;  Score 771.5;  DB 1;  Length 173;
  Best Local Similarity   78.2%;  Pred. No. 2.3e-63;
  Matches  136;  Conservative   18;  Mismatches   19;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGHQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              |||||||||||:|||||||||| |||: |||:||::| :| | |:::||  :|||||:|
Db         61 DYPDYQQWMGFSDSIRSCRLIP-HTGSHRMRLYEKEDHKGVMMELSEDCSCIQDRFHLSE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              | ||:|||| |||||||:|||||||||| ||||| ||||::|| ||||||:| |
Db        120 VRSLHVLEGCWVLYEMPNYRGRQYLLRPQEYRRYHDWGAVDAKAGSLRRVVDLY 173


RESULT 15
CRGA_RAT
ID   CRGA_RAT       STANDARD;      PRT;   173 AA.
AC   P10065;
DT   01-MAR-1989 (Rel. 10, Created)
DT   01-AUG-1991 (Rel. 19, Last sequence update)
DT   01-OCT-2004 (Rel. 45, Last annotation update)
DE   Gamma crystallin A (Gamma crystallin 1-1).
GN   Name=Cryga;
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90236314; PubMed=1970548;
RA   den Dunnen J.T., van Neck J.W., Cremers F.P.M., Lubsen N.H.,
RA   Schoenmakers J.G.G.;
RT   "Isolation and characterization of the rat glutamine synthetase-
RT   encoding gene.";
RL   Gene 87:225-232(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=87060933; PubMed=3783678;
RA   den Dunnen J.T., Moormann R.J.M., Lubsen N.H., Schoenmakers J.G.G.;
RT   "Concerted and divergent evolution within the rat gamma-crystallin
RT   gene family.";
RL   J. Mol. Biol. 189:37-46(1986).
CC   -!- FUNCTION: Crystallins are the dominant structural components of
CC       the vertebrate eye lens.
CC   -!- DOMAIN: Has a two-domain beta-structure, folded into four very
CC       similar Greek key motifs.
CC   -!- MISCELLANEOUS: There are six different gamma crystallins
CC       identified in rat lens.
CC   -!- SIMILARITY: Belongs to the beta/gamma-crystallin family.
CC   -!- SIMILARITY: Contains 4 beta/gamma crystallin 'Greek key' domains.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M19359; AAA40981.1; -.
DR   PIR; A24060; A24060.
DR   HSSP; P02526; 1AMM.
DR   RGD; 2419; Cryga.
DR   InterPro; IPR001064; Crystallin.
DR   InterPro; IPR011024; G_crystallin_SF.
DR   Pfam; PF00030; Crystall; 2.
DR   PRINTS; PR01367; BGCRYSTALLIN.
DR   SMART; SM00247; XTALbg; 2.
DR   PROSITE; PS50915; CRYSTALLIN_BETAGAMMA; 4.
KW   Eye lens protein; Multigene family; Repeat.
FT   INIT_MET      0      0
FT   DOMAIN        1     39       Beta/gamma crystallin 'Greek key' 1.
FT   DOMAIN       40     82       Beta/gamma crystallin 'Greek key' 2.
FT   DOMAIN       83     86       Connecting peptide.
FT   DOMAIN       87    127       Beta/gamma crystallin 'Greek key' 3.
FT   DOMAIN      128    170       Beta/gamma crystallin 'Greek key' 4.
FT   CONFLICT     15     15       C -> S (in Ref. 2).
SQ   SEQUENCE   173 AA;  20962 MW;  A561FCFCA70F8620 CRC64;

  Query Match             69.4%;  Score 769.5;  DB 1;  Length 173;
  Best Local Similarity   77.6%;  Pred. No. 3.5e-63;
  Matches  135;  Conservative   20;  Mismatches   18;  Indels    1;  Gaps    1;

Qy          2 GRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |:| | |||||||  | |:||||||| |||||||||| ||||||||||||||:|||||||
Db          1 GKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSIRVDSGCWMLYERPNYQGYQYFLRRG 60

Qy         62 DYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLTE 121
              ||||||||||:||||||| || :| : |:|:|||||:|| :||:|:||  : ||| | |
Db         61 DYPDYQQWMGFSDSIRSCRSIP-YTSSHRIRLYERDDYRGLVSELTEDCSCIHDRFRLNE 119

Qy        122 VHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMDFY 175
              ::|::|||||||||||:||||||||||:|||| |||||:|||||||||| |
Db        120 IYSMHVLEGSWVLYEMPNYRGRQYLLRPGDYRRYHDWGAMDAKVGSLRRVMDLY 173


Search completed: January 14, 2005, 12:10:32
Job time : 196 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 12:03:35 ; Search time 21 Seconds
(without alignments)
622.126 Million cell updates/sec

Title: US-10-030-605A-21
Perfect score: 1108
Sequence: 1 MGRIKFKEDRGFQGHYYSCN.....PNSSSVDKLAAALEHHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	300	27.1	205	2	US-08-729-152-8	Sequence 8, Appli
2	281.5	25.4	215	2	US-08-729-152-1	Sequence 1, Appli
3	100	9.0	42	1	US-08-664-449-39	Sequence 39, Appl
4	86	7.8	378	4	US-08-979-847B-122	Sequence 122, App
5	86	7.8	398	4	US-08-979-847B-121	Sequence 121, App
6	82	7.4	320	2	US-08-757-653-163	Sequence 163, App
7	82	7.4	320	2	US-08-823-516-61	Sequence 61, Appl
8	82	7.4	320	3	US-08-759-038-102	Sequence 102, App
9	82	7.4	320	3	US-08-758-314-102	Sequence 102, App
10	82	7.4	320	4	US-09-684-938-102	Sequence 102, App
11	82	7.4	320	4	US-09-308-825A-102	Sequence 102, App
12	82	7.4	320	4	US-09-940-244-61	Sequence 61, Appl
13	80.5	7.3	655	4	US-09-270-767-43477	Sequence 43477, A
14	79.5	7.2	139	4	US-09-621-976-5088	Sequence 5088, Ap
15	79.5	7.2	300	3	US-09-561-756-36	Sequence 36, Appl
16	79.5	7.2	300	3	US-09-227-721-36	Sequence 36, Appl
17	79.5	7.2	300	4	US-09-954-697-36	Sequence 36, Appl
18	79.5	7.2	315	2	US-08-484-956-91	Sequence 91, Appl
19	79.5	7.2	315	2	US-08-757-653-91	Sequence 91, Appl
20	79.5	7.2	315	3	US-08-520-946-91	Sequence 91, Appl
21	79.5	7.2	315	4	US-09-655-378A-91	Sequence 91, Appl
22	79.5	7.2	1010	4	US-09-654-449-2	Sequence 2, Appli
23	78.5	7.1	243	4	US-09-543-681A-7356	Sequence 7356, Ap
24	77	6.9	225	4	US-09-456-090A-94	Sequence 94, Appl
25	77	6.9	225	4	US-09-453-234-94	Sequence 94, Appl
26	76	6.9	229	4	US-09-546-043-8	Sequence 8, Appli
27	76	6.9	338	4	US-09-546-043-4	Sequence 4, Appli
28	76	6.9	546	3	US-09-457-040B-24	Sequence 24, Appl
29	76	6.9	546	4	US-09-538-092-668	Sequence 668, App
30	75.5	6.8	323	3	US-08-816-977-21	Sequence 21, Appl
31	75.5	6.8	323	4	US-09-334-477-21	Sequence 21, Appl
32	75	6.8	225	4	US-09-456-090A-58	Sequence 58, Appl
33	75	6.8	225	4	US-09-453-234-58	Sequence 58, Appl
34	75	6.8	454	4	US-09-450-209-16	Sequence 16, Appl
35	75	6.8	738	3	US-08-989-385-1	Sequence 1, Appli
36	75	6.8	738	4	US-09-593-826-1	Sequence 1, Appli
37	74.5	6.7	316	4	US-09-387-375-9	Sequence 9, Appli
38	74.5	6.7	316	4	US-10-041-400A-9	Sequence 9, Appli
39	74	6.7	225	4	US-09-456-090A-54	Sequence 54, Appl
40	74	6.7	225	4	US-09-456-090A-56	Sequence 56, Appl
41	74	6.7	225	4	US-09-456-090A-60	Sequence 60, Appl
42	74	6.7	225	4	US-09-456-090A-62	Sequence 62, Appl
43	74	6.7	225	4	US-09-456-090A-66	Sequence 66, Appl
44	74	6.7	225	4	US-09-456-090A-68	Sequence 68, Appl
45	74	6.7	225	4	US-09-456-090A-70	Sequence 70, Appl

ALIGNMENTS

```
RESULT 1
US-08-729-152-8
; Sequence 8, Application US/08729152
; Patent No. 5871739
;  GENERAL INFORMATION:
;    APPLICANT:  Inoue, Eri
;    TITLE OF INVENTION:  Pharmaceutical Composition
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  Southern Building, Suite 700, 805 Fifteenth
;      STREET:  Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/729,152
;      FILING DATE:  11-OCT-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 291993/1995
;      FILING DATE:  13-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warren M. Cheek, Jr.
;      REGISTRATION NUMBER:  33,367
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-8850
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  8:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  205 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  protein
US-08-729-152-8

  Query Match           27.1%;  Score 300;  DB 2;  Length 205;
  Best Local Similarity  37.1%;  Pred. No. 2.4e-24;
  Matches   63;  Conservative   31;  Mismatches   72;  Indels    4;  Gaps    3;
```

```
Qy          8 EDRGFQGHYYSCNSDCPNL-QPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGDYPDY 66
              |    |||| :  |  |||| :    :  |: | :| |: ||: | :| |:   :|:|| :
Db         23 EQENFQGHSHELNGPCPNLKETGVEKAGSVLVQAGPWVGYEQANCKGEQFVFEKGEYPRW 82

Qy         67 QQWMGF--NDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSE-ITDDCPSLQDRFHLTEVH 123
                |      ||: | | |  :  :: :||  :| |:  | | || ||     :  :|
Db         83 DSWTSSRRTDSLSSLRPIKVDSQEHKIILYENPNFTGKKMEIIDDDVPSFHAHGYQEKVS 142

Qy        124 SLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVMD 173
              |: |  |:|| |: | ||| |||| |:|:   |:|| : :| |:||: |
Db        143 SVRVQSGTWVGYQYPGYRGLQYLLEKGDYKDSSDFGAPHPQVQSVRRIRD 192


RESULT 2
US-08-729-152-1
; Sequence 1, Application US/08729152
; Patent No. 5871739
;  GENERAL INFORMATION:
;    APPLICANT:  Inoue, Eri
;    TITLE OF INVENTION:  Pharmaceutical Composition
;    NUMBER OF SEQUENCES:  47
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Wenderoth, Lind & Ponack
;      STREET:  Southern Building, Suite 700, 805 Fifteenth
;      STREET:  Street, N.W.
;      CITY:  Washington
;      STATE:  D.C.
;      COUNTRY:  USA
;      ZIP:  20005
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Diskette - 3.5 inch, 1.44 Mb storage
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/729,152
;      FILING DATE:  11-OCT-1996
;      CLASSIFICATION:  424
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  JP 291993/1995
;      FILING DATE:  13-OCT-1995
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Warren M. Cheek, Jr.
;      REGISTRATION NUMBER:  33,367
;      REFERENCE/DOCKET NUMBER:
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  202-371-8850
;      TELEFAX:
;      TELEX:
;  INFORMATION FOR SEQ ID NO:  1:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  215 amino acids
;      TYPE:  amino acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:
;      DESCRIPTION:  protein
US-08-729-152-1

  Query Match             25.4%;  Score 281.5;  DB 2;  Length 215;
  Best Local Similarity   34.1%;  Pred. No. 2.5e-22;
  Matches   62;  Conservative   37;  Mismatches   70;  Indels   13;  Gaps    6;

Qy          3 RIKFKEDRGFQGHYYSCNSDCPNL-QPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              :|   :   |||      | |||: :  |    |::| || |: ||  :: | |: | ||
Db         32 KITIYDQENFQGKRMEFTSSCPNVSERSFDNVRSLKVESGAWIGYEHTSFCGQQFILERG 91

Qy         62 DYPDYQQWMGFN----DSIRSCR-LIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDR 116
                :|| :  | | |    : : | |      :    :| |:|:::| |:  ||:|| ||||
Db         92 EYPRWDAWSGSNAYHMERLMSFRPFCSANHKESKMTIFEKENFIGRQWEISDDYPSLQAM 151

Qy        117 -FHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLR----PGEYRRYLDWG--AMNAKVGSLR 169
               :  || |: :  |:|| |    ||| ||:|:     |:|: : :||  |  ::: |:|
Db        152 GWFNNEVGSMKIQSGAWVCYHYLGYRGYQYILKCDHHEGDYKHWREWGSHAQTSQIQSIR 211

Qy        170 RV 171
               |:
Db        212 RI 213


RESULT 3
US-08-664-449-39
; Sequence 39, Application US/08664449
; Patent No. 5766905
;  GENERAL INFORMATION:
;    APPLICANT:  Studier, F. W.
;    APPLICANT:  Rosenberg, Alan H.
;    TITLE OF INVENTION:  Cytoplasmic Bacteriophage Display System
;    NUMBER OF SEQUENCES:  70
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Brookhaven National Laboratory
;      STREET:  Building 902C
;      CITY:  Upton
;      STATE:  NY
;      COUNTRY:  US
;      ZIP:  11973
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/664,449
;      FILING DATE:  17-June-1996
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bogosian, Margaret
;      REGISTRATION NUMBER:  25,324
;      REFERENCE/DOCKET NUMBER:  AUI-9618
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (516) 344-7338
;      TELEFAX:  (516) 344-3729
;  INFORMATION FOR SEQ ID NO:  39:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  42 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-664-449-39

  Query Match             9.0%;  Score 100;  DB 1;  Length 42;
  Best Local Similarity   53.8%;  Pred. No. 0.00095;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels   18;  Gaps    1;

Qy        177 DPNSSSVDKLAAAL-----------------EHHHHHH 197
              |||||||||||||||                 |||||||
Db          4 DPNSSSVDKLAAALEIKRASQPELAPEDPEDVEHHHHHH 42


RESULT 4
US-08-979-847B-122
; Sequence 122, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
```

```
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 122:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 378 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 122:
US-08-979-847B-122

  Query Match             7.8%;  Score 86;  DB 4;  Length 378;
  Best Local Similarity   78.9%;  Pred. No. 0.65;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        179 NSSSVDKLAAALEHHHHHH 197
              |  |: ||||||||||||
Db        360 NFKSLPKLAAALEHHHHHH 378


RESULT 5
US-08-979-847B-121
; Sequence 121, Application US/08979847B
; Patent No. 6582703
;   GENERAL INFORMATION:
;        APPLICANT: PERRON, HERVE
;                   BESEME, FREDERIC
;                   BEDIN, FREDERIC
;                   PARANHOS-BACCALA, GLAUCIA
;                   KOMURIAN-PRADEL, FLORENCE
;                   JOLIVET-REYNAUD, COLETTE
;                   MANDRAND, BERNARD
;                   GARSON, JEREMY
;                   TUKE, PHILIP
;        TITLE OF INVENTION: VIRAL MATERIAL AND NUCLEOTIDE FRAGMENTS
;                            ASSOCIATED WITH MULTIPLE SCLEROSIS, FOR DIAGNOSTIC, PROPHYL
;                            THERAPEUTIC PURPOSES
;        NUMBER OF SEQUENCES: 210
;        CORRESPONDENCE ADDRESS:
;             ADDRESSEE: OLIFF & BERRIDGE, PLC
;             STREET: P.O. BOX 19928
;             CITY: ALEXANDRIA
;             STATE: VA
;             COUNTRY: USA
;             ZIP: 22320
;        COMPUTER READABLE FORM:
;             MEDIUM TYPE: Floppy disk
;             COMPUTER: IBM PC compatible
;             OPERATING SYSTEM: PC-DOS/MS-DOS
;             SOFTWARE: PatentIn Release #1.0, Version #1.30
;        CURRENT APPLICATION DATA:
;             APPLICATION NUMBER: US/08/979,847B
;             FILING DATE: 26-No. 6582703-1997
;             CLASSIFICATION: <Unknown>
;        ATTORNEY/AGENT INFORMATION:
;             NAME: BERRIDGE, WILLIAM P.
;             REGISTRATION NUMBER: 30,024
;             REFERENCE/DOCKET NUMBER: WPB 39046A
;        TELECOMMUNICATION INFORMATION:
;             TELEPHONE: 703-836-6400
;             TELEFAX: 703-836-2787
;   INFORMATION FOR SEQ ID NO: 121:
;        SEQUENCE CHARACTERISTICS:
;             LENGTH: 398 amino acids
;             TYPE: amino acid
;             STRANDEDNESS: single
;             TOPOLOGY: linear
;        MOLECULE TYPE: peptide
;        SEQUENCE DESCRIPTION: SEQ ID NO: 121:
US-08-979-847B-121

  Query Match             7.8%;  Score 86;  DB 4;  Length 398;
  Best Local Similarity   78.9%;  Pred. No. 0.69;
  Matches   15;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        179 NSSSVDKLAAALEHHHHHH 197
              |  |: ||||||||||||
Db        380 NFKSLPKLAAALEHHHHHH 398


RESULT 6
US-08-757-653-163
; Sequence 163, Application US/08757653
; Patent No. 5843669
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
;    TITLE OF INVENTION:  Cleavage Of Nucleic Acid Using
;    TITLE OF INVENTION:  Thermostable FEN-1 Endonucleases
;    NUMBER OF SEQUENCES:  190
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/757,653
;      FILING DATE:
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02565
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  163:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  320 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-757-653-163

  Query Match             7.4%;  Score 82;  DB 2;  Length 320;
```

```
  Best Local Similarity   19.2%;  Pred. No. 1.4;
  Matches   43;  Conservative   24;  Mismatches   51;  Indels  106;  Gaps    9;

Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE----------ITDD---CP 111
              |  |  |   || :     |  |: :  || ||    :::          :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
               | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 ----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------- 186
                      :||:: | : :| |:     :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276

Qy        187 --------------------------------AAALEHHHHHH 197
                                               |||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320


RESULT 7
US-08-823-516-61
; Sequence 61, Application US/08823516
; Patent No. 5994069
;  GENERAL INFORMATION:
;    APPLICANT:  Hall, Jeff G.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Mast, Andrea L.
;    APPLICANT:  Brow, Mary Ann D.
;    TITLE OF INVENTION:  Detection Of Nucleic Acids By Multiple
;    TITLE OF INVENTION:  Sequential Invasive Cleavages
;    NUMBER OF SEQUENCES:  163
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/823,516
;      FILING DATE:  24-MAR-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US97/01072
;      FILING DATE:  21-JAN-1997
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/759,038
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/758,314
;      FILING DATE:  02-DEC-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/756,386
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02736
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  61:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  320 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-823-516-61

  Query Match             7.4%;  Score 82;  DB 2;  Length 320;
  Best Local Similarity   19.2%;  Pred. No. 1.4;
  Matches   43;  Conservative   24;  Mismatches   51;  Indels  106;  Gaps    9;

Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE----------ITDD---CP 111
              |  |  |   || :     |  |: :  || ||    :::          :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
               | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 ----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------- 186
                      :||:: | : :| |:     :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276

Qy        187 --------------------------------AAALEHHHHHH 197
                                               |||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320


RESULT 8
US-08-759-038-102
; Sequence 102, Application US/08759038
; Patent No. 6090543
;  GENERAL INFORMATION:
;    APPLICANT:  Prudent, James R.
;    APPLICANT:  Hall, Jeff G.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Brow, Mary Ann D.
;    APPLICANT:  Dahlberg, James E.
;    TITLE OF INVENTION:  Cleavage Of Nucleic Acids
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/759,038
;      FILING DATE:  02-DEC-1996
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02574
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
```

```
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  320 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-759-038-102

  Query Match             7.4%;  Score 82;  DB 3;  Length 320;
  Best Local Similarity   19.2%;  Pred. No. 1.4;
  Matches   43;  Conservative   24;  Mismatches   51;  Indels  106;  Gaps    9;

Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE-----------ITDD---CP 111
              |  |  |   || :      |  |: :  || ||    :::           :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
              | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 -----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------ 186
                        :||:: | : :| |:     :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276

Qy        187 --------------------------------AAALEHHHHHH 197
                                              ||||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320


RESULT 9
US-08-758-314-102
; Sequence 102, Application US/08758314
; Patent No. 6090606
;  GENERAL INFORMATION:
;    APPLICANT:  Kaiser, Michael W.
;    APPLICANT:  Lyamichev, Victor I.
;    APPLICANT:  Lyamichev, Natasha
;    TITLE OF INVENTION:  Improved Cleavage Agents
;    NUMBER OF SEQUENCES:  134
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Medlen & Carroll, LLP
;      STREET:  220 Montgomery Street, Suite 2200
;      CITY:  San Francisco
;      STATE:  California
;      COUNTRY:  United States Of America
;      ZIP:  94104
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.30
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/758,314
;      FILING DATE:  02-DEC-1996
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/
;      FILING DATE:  29-NOV-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/682,853
;      FILING DATE:  12-JUL-1996
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/599,491
;      FILING DATE:  24-JAN-1996
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Ingolia, Diane E.
;      REGISTRATION NUMBER:  40,027
;      REFERENCE/DOCKET NUMBER:  FORS-02575
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 705-8410
;      TELEFAX:  (415) 397-8338
;  INFORMATION FOR SEQ ID NO:  102:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  320 amino acids
;      TYPE:  amino acid
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  protein
US-08-758-314-102

  Query Match             7.4%;  Score 82;  DB 3;  Length 320;
  Best Local Similarity   19.2%;  Pred. No. 1.4;
  Matches   43;  Conservative   24;  Mismatches   51;  Indels  106;  Gaps    9;

Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE-----------ITDD---CP 111
              |  |  |   || :      |  |: :  || ||    :::           :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
              | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 -----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------ 186
                        :||:: | : :| |:     :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276

Qy        187 --------------------------------AAALEHHHHHH 197
                                              ||||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320


RESULT 10
US-09-684-938-102
; Sequence 102, Application US/09684938
; Patent No. 6555357
; GENERAL INFORMATION:
;  APPLICANT: Kaiser, Michael W.
;  APPLICANT:  Lyamichev, Victor I.
;  APPLICANT:  Lyamichev, Natasha
;  TITLE OF INVENTION: Improved Cleavage Agents
;  FILE REFERENCE: FORS-03755
;  CURRENT APPLICATION NUMBER: US/09/684,938
;  CURRENT FILING DATE:  2000-10-06
;  PRIOR APPLICATION NUMBER: 09/308,825
;  PRIOR FILING DATE: 1999-05-25
;  PRIOR APPLICATION NUMBER: 08/757,653
;  PRIOR FILING DATE: 1996-11-29
;  PRIOR APPLICATION NUMBER: 08/758,314
;  PRIOR FILING DATE: 1996-12-02
;  PRIOR APPLICATION NUMBER: PCT/US97/21783
;  PRIOR FILING DATE: 1997-11-29
;  NUMBER OF SEQ ID NOS: 188
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
;   LENGTH: 320
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:  Synthetic
US-09-684-938-102

  Query Match             7.4%;  Score 82;  DB 4;  Length 320;
  Best Local Similarity   19.2%;  Pred. No. 1.4;
  Matches   43;  Conservative   24;  Mismatches   51;  Indels  106;  Gaps    9;

Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE-----------ITDD---CP 111
              |  |  |   || :      |  |: :  || ||    :::           :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
              | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 -----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------ 186
                        :||:: | : :| |:     :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276
```

```
Qy        187 --------------------------------AAALEHHHHHH 197
                                               ||||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320
```

RESULT 11
US-09-308-825A-102
; Sequence 102, Application US/09308825A
; Patent No. 6562611
; GENERAL INFORMATION:
; APPLICANT: Kaiser, Michael W.
; APPLICANT: Lyamichev, Victor I.
; APPLICANT: Lyamichev, Natasha
; TITLE OF INVENTION: Improved Cleavage Agents
; FILE REFERENCE: FORS-03755
; CURRENT APPLICATION NUMBER: US/09/308,825A
; CURRENT FILING DATE: 1999-10-08
; PRIOR APPLICATION NUMBER: 08/757,653
; PRIOR FILING DATE: 1996-11-29
; PRIOR APPLICATION NUMBER: 08/758,314
; PRIOR FILING DATE: 1996-12-02
; PRIOR APPLICATION NUMBER: PCT/US97/21783
; PRIOR FILING DATE: 1997-11-29
; NUMBER OF SEQ ID NOS: 188
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 102
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-308-825A-102

Query Match 7.4%; Score 82; DB 4; Length 320;
Best Local Similarity 19.2%; Pred. No. 1.4;
Matches 43; Conservative 24; Mismatches 51; Indels 106; Gaps 9;

```
Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE-----------ITDD---CP 111
                |  |  |   || :     |  |: :  || ||    :::           :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
                | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 -----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------ 186
                         :||:: | : :| |:      :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276

Qy        187 --------------------------------AAALEHHHHHH 197
                                               ||||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320
```

RESULT 12
US-09-940-244-61
; Sequence 61, Application US/09940244
; Patent No. 6692917
; GENERAL INFORMATION:
; APPLICANT: Neri, Bruce P.
; APPLICANT: Hall, Jeff G.
; APPLICANT: Lyamichev, Victor
; APPLICANT: Smith, Lloyd M.
; TITLE OF INVENTION: Reactions on Dendrimers
; FILE REFERENCE: FORS-06478
; CURRENT APPLICATION NUMBER: US/09/940,244
; CURRENT FILING DATE: 2002-05-06
; NUMBER OF SEQ ID NOS: 422
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-09-940-244-61

Query Match 7.4%; Score 82; DB 4; Length 320;
Best Local Similarity 19.2%; Pred. No. 1.4;
Matches 43; Conservative 24; Mismatches 51; Indels 106; Gaps 9;

```
Qy         71 GFNDSIRSCRLIPQHT---GTFRMRI--YERDDFRGQMSE-----------ITDD---CP 111
                |  |  |   || :     |  |: :  || ||    :::           :| |
Db        106 GSEDFPRQLALIKELVDLLGLARLEVPGYEADDVLASLAKKAEKEGYEVRILTADKDLYQ 165

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYL--------------- 156
                | || |:  :|     ||  :    |::   :| ||| ::  |
Db        166 LLSDRIHV--LHP----EGYLI---TPAWLWEKYGLRPDQWADYRALTGDESDNLPGVKG 216

Qy        157 -----------DWGAMNAKVGSLRRVMDFYSDPNSSSVDKL------------------ 186
                         :||:: | : :| |:      :   : :| |
Db        217 IGEKTARKLLEEWGSLEALLKNLDRLKPAIREKILAHMDDLKLSWDLAKVRTDLPLEVDF 276

Qy        187 --------------------------------AAALEHHHHHH 197
                                               ||||||||||
Db        277 AKRREPDRERLRAFLERLEFGSLLHEFGLLESPKAALEHHHHHH 320
```

RESULT 13
US-09-270-767-43477
; Sequence 43477, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43477
; LENGTH: 655
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43477

Query Match 7.3%; Score 80.5; DB 4; Length 655;
Best Local Similarity 20.1%; Pred. No. 5.4;
Matches 42; Conservative 34; Mismatches 102; Indels 31; Gaps 8;

```
Qy          3 RIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRV--LSGCWMLYERPNYQGHQYFLR- 59
              ::||  :: || |       |     |   :  : |   :|  : | :  |    |  |:
Db        180 QLKFNGEKMFQVH------DIETFDPQHVQLKIVPVEKTAGLHVEYSKQEYGKSQLELQP 233

Qy         60 -------RGDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYE-RDDFRGQMSEITDDCP 111
                    :||       : |: ::   |:: ||   ||:  |   ::  |  : :| |
Db        234 TFGVPYTKGDI------VIFHITLPLERMMEQH---FRLECYSMSNELLGSATLVTSDLT 284

Qy        112 SLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRP----GEYRRYLDWGAMNAKVGS 167
                :   ||  : |    : :     :|    : |   |       |  |       |  ||
Db        285 GSEGVLHL-PIKSAKNADETLARLRLPYVAVQPYRYSPLDFKNTYAHYWPKSWPNLDVGH 343

Qy        168 LRRVMDFYSDPNSSSVDKLAAALEHHHHH 196
                  : :|  :   : :|: |  | ||
Db        344 RGNGKSYIADAPAERENTIASFLSAHEHH 372
```

RESULT 14
US-09-621-976-5088
; Sequence 5088, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.

```
;  APPLICANT:  Giordano, J.Y.
;  TITLE OF INVENTION: ESTs and Encoded Human Proteins.
;  FILE REFERENCE: GENSET.054PR2
;  CURRENT APPLICATION NUMBER: US/09/621,976
;  CURRENT FILING DATE:  2000-07-21
;  NUMBER OF SEQ ID NOS: 19335
;  SOFTWARE: Patent.pm
; SEQ ID NO 5088
;   LENGTH: 139
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: UNSURE
;   LOCATION: 104
;   OTHER INFORMATION: Xaa = Ala,Asp
;   NAME/KEY: UNSURE
;   LOCATION: 114
;   OTHER INFORMATION: Xaa = Ile,Ser
US-09-621-976-5088

  Query Match             7.2%;  Score 79.5;  DB 4;  Length 139;
  Best Local Similarity   41.7%;  Pred. No. 0.81;
  Matches   20;  Conservative    7;  Mismatches   16;  Indels    5;  Gaps    2;

Qy        124 SLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMN--AKVGSLR 169
              |: |  | |: ||  :: ||| |||| |     :| | :    :|||| 
Db         45 SVWVKSGLWIAYEGSNFLGRQILLRPNEIP---NWTAFSRWKTIGSLR 89


RESULT 15
US-09-561-756-36
; Sequence 36, Application US/09561756
; Patent No. 6376226
; GENERAL INFORMATION:
;  APPLICANT: Alnemri, Emad S.
;  TITLE OF INVENTION: RECOMBINANT, ACTIVE CASPASES AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 480140.431
;  CURRENT APPLICATION NUMBER: US/09/561,756
;  CURRENT FILING DATE:  2000-04-26
;  PRIOR APPLICATION NUMBER: 09/227,721
;  PRIOR FILING DATE: 1999-01-08
;  NUMBER OF SEQ ID NOS: 116
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 36
;   LENGTH: 300
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Rev-Caspase-6 constructed from human caspase-6
US-09-561-756-36

  Query Match             7.2%;  Score 79.5;  DB 3;  Length 300;
  Best Local Similarity   31.5%;  Pred. No. 2.3;
  Matches   28;  Conservative   10;  Mismatches   22;  Indels   29;  Gaps    6;

Qy        128 LEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAM----NAKVGSLRRVMDFYSDPN---- 179
              : ||| : ::    |:        |   |::  :    | || | ||| ||  ||:     
Db         35 VNGSWYIQDLCEMLGK--------YGSSLEFTELLTLVNRKV-SQRRV-DFCKDPSAIGK 84

Qy        180 -------SSSVDKL----AAALEHHHHHH 197
                    |    ||     : |||||||
Db         85 KQVPCFASMLTKKLHFFPKSNLEHHHHHH 113


Search completed: January 14, 2005, 12:06:50
Job time : 23 secs
```

GenCore version 5.1.6

OM protein - protein search, using sw model

Run on: January 14, 2005, 12:03:39 ; Search time 145 Seconds
(without alignments)
490.855 Million cell updates/sec

Title: US-10-030-605A-21
Perfect score: 1108
Sequence: 1 MGRIKFKEDRGFQGHYYSCN.....PNSSSVDKLAAALEHHHHHH 197

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1608061 seqs, 361289386 residues

Total number of hits satisfying chosen parameters: 1608061

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	424.5	38.3	89	17	US-10-425-115-346968	Sequence 346968,
2	320	28.9	616	15	US-10-108-260A-3095	Sequence 3095, Ap
3	320	28.9	781	10	US-09-866-050A-668	Sequence 668, App
4	257	23.2	197	17	US-10-723-860-1946	Sequence 1946, Ap
5	253	22.8	97	9	US-09-925-298-655	Sequence 655, App
6	253	22.8	97	14	US-10-102-806-655	Sequence 655, App
7	210.5	19.0	511	14	US-10-104-047-2339	Sequence 2339, Ap
8	200	18.1	138	14	US-10-106-698-5959	Sequence 5959, Ap
9	163	14.7	83	9	US-09-864-761-44293	Sequence 44293, A
10	147	13.3	30	9	US-09-808-602-62	Sequence 62, Appl
11	123.5	11.1	826	15	US-10-395-241-18	Sequence 18, Appl
12	123	11.1	168	14	US-10-360-053-2	Sequence 2, Appli
13	123	11.1	168	14	US-10-360-053-4	Sequence 4, Appli
14	119	10.7	21	14	US-10-010-160-68	Sequence 68, Appl
15	100.5	9.1	151	14	US-10-360-053-10	Sequence 10, Appl
16	100.5	9.1	159	14	US-10-360-053-12	Sequence 12, Appl
17	95	8.6	526	9	US-09-731-221-79	Sequence 79, Appl
18	93.5	8.4	381	14	US-10-141-531-47	Sequence 47, Appl
19	93.5	8.4	381	14	US-10-290-072-47	Sequence 47, Appl
20	93	8.4	516	9	US-09-804-626-4	Sequence 4, Appli
21	93	8.4	519	14	US-10-359-369-34	Sequence 34, Appl
22	92	8.3	381	14	US-10-141-531-48	Sequence 48, Appl
23	92	8.3	381	14	US-10-141-531-49	Sequence 49, Appl
24	92	8.3	381	14	US-10-141-531-50	Sequence 50, Appl
25	92	8.3	381	14	US-10-141-531-51	Sequence 51, Appl
26	92	8.3	381	14	US-10-141-531-52	Sequence 52, Appl
27	92	8.3	381	14	US-10-141-531-53	Sequence 53, Appl
28	92	8.3	381	14	US-10-141-531-54	Sequence 54, Appl
29	92	8.3	381	14	US-10-141-531-56	Sequence 56, Appl
30	92	8.3	381	14	US-10-141-531-58	Sequence 58, Appl
31	92	8.3	381	14	US-10-290-072-48	Sequence 48, Appl
32	92	8.3	381	14	US-10-290-072-49	Sequence 49, Appl
33	92	8.3	381	14	US-10-290-072-50	Sequence 50, Appl
34	92	8.3	381	14	US-10-290-072-51	Sequence 51, Appl
35	92	8.3	381	14	US-10-290-072-52	Sequence 52, Appl
36	92	8.3	381	14	US-10-290-072-53	Sequence 53, Appl
37	92	8.3	381	14	US-10-290-072-54	Sequence 54, Appl
38	92	8.3	381	14	US-10-290-072-56	Sequence 56, Appl
39	92	8.3	381	14	US-10-290-072-58	Sequence 58, Appl
40	92	8.3	481	14	US-10-141-531-59	Sequence 59, Appl
41	92	8.3	481	14	US-10-290-072-59	Sequence 59, Appl
42	90.5	8.2	653	14	US-10-369-493-5789	Sequence 5789, Ap
43	89	8.0	556	10	US-09-819-266-28	Sequence 28, Appl
44	89	8.0	559	10	US-09-819-266-25	Sequence 25, Appl
45	88	7.9	518	9	US-09-804-626-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-10-425-115-346968
; Sequence 346968, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 346968
; LENGTH: 89
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_79599C.1.pep
US-10-425-115-346968

Query Match 38.3%; Score 424.5; DB 17; Length 89;
Best Local Similarity 82.2%; Pred. No. 2.7e-36;
Matches 74; Conservative 6; Mismatches 9; Indels 1; Gaps 1;

```
Qy          1 MGRIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              ||:| | ||||||| | |:||||||| |||||||:|| ||||||||||||||||||||
Db          1 MGKITFYEDRGFQGRCYECSSDCPNLQTYFSRCNSVRVDSGCWMLYERPNYQGHQYFLRR 60

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFR 90
              |||||||||:||:| |||||||| |||: |
Db         61 GDYPDYQQWLGFSDYIRSCRLIP-HTGSHR 89
```

```
RESULT 2
US-10-108-260A-3095
; Sequence 3095, Application US/10108260A
; Publication No. US20040005560A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20040005560A1el full length cDNA
;  FILE REFERENCE: H1-A0106
;  CURRENT APPLICATION NUMBER: US/10/108,260A
;  CURRENT FILING DATE:  2002-03-27
;  NUMBER OF SEQ ID NOS: 5458
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3095
;   LENGTH: 616
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-108-260A-3095

  Query Match             28.9%;  Score 320;  DB 15;  Length 616;
  Best Local Similarity   33.0%;  Pred. No. 2.4e-24;
  Matches   68;  Conservative   34;  Mismatches   78;  Indels   26;  Gaps    2;

Qy          8 EDRGFQGHYYSCNSDCPNLQ-------PYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              |  ||||  :  : |  |||        |: :   |:||| |||: ||:  ::|||| |
Db        129 EAPGFQGRSWEVSRDIYNLQQPEDSQSPHLASVGSLRVLGGCWVGYEKEGFRGHQYLLEE 188

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |:|||:  | |::: : | |:|    |   : ::|  || |   |::   | ::   |
Db        189 GEYPDWSHWGGYDELLTSLRVIRTDFGDPAVVLFEAMDFEGHGVEVSKALPDVELVQHGP 248

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRVM-------- 172
              :::|| | || |:   : | ||:|  | ||   |||| |: : ||: |:
Db       -249 STQAIHVLSGVWVAYQEVGFSGEQYVLEKGVYRNCEDWGAGNSTLASLQPVLQVGEHDLH 308

Qy        173 ----------DFYSDPNSSSVDKLA 187
                       ||  |  |   |: |
Db        309 FVSKIQLFSRPDFLGDHFSFEDDQAA 334


RESULT 3
US-09-866-050A-668
; Sequence 668, Application US/09866050A
; Publication No. US20030040471A1
; GENERAL INFORMATION:
;  APPLICANT: Watson, James D.
;  APPLICANT:  Strachan, Lorna
;  APPLICANT:  Sleeman, Matthew
;  APPLICANT:  Onrust, Rene
;  APPLICANT:  Murison, James G.
;  APPLICANT:  Kumble, Krishanand D.
;  TITLE OF INVENTION: Compositions Isolated From Skin Cells
;  TITLE OF INVENTION:  and Methods for Their Use
;  FILE REFERENCE: 11000.1011c4U
;  CURRENT APPLICATION NUMBER: US/09/866,050A
;  CURRENT FILING DATE:  2001-05-24
;  NUMBER OF SEQ ID NOS: 725
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 668
;   LENGTH: 781
;   TYPE: PRT
;   ORGANISM: Mouse
US-09-866-050A-668

  Query Match             28.9%;  Score 320;  DB 10;  Length 781;
  Best Local Similarity   33.3%;  Pred. No. 3.2e-24;
  Matches   64;  Conservative   34;  Mismatches   74;  Indels   20;  Gaps    2;

Qy          8 EDRGFQGHYYSCNSDCPNLQ-------PYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRR 60
              |  ||||  :  : |  |||        |  :   |:|:| |||: ||:  ::|||| |
Db        294 EAPGFQGQSWEVSGDIYNLQQPEDSQSPQLTSVGSLRILGGCWVGYEKEGFRGHQYLLEE 353

Qy         61 GDYPDYQQWMGFNDSIRSCRLIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDRFHLT 120
              |:| |:  | |::: : | |:|    |   : ::|  ||:|   |::   | ::   |
Db        354 GEYADWSHWGGYDELLTSLRVIRTDFGDPAVVLFEDMDFQGHRVEVSSALPDVELAQHGP 413

Qy        121 EVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEYRRYLDWGAMNAKVGSLRRV--------- 171
              :::|| | || ||   : | ||:|  | ||   |||: |  :|||: |
Db        414 STQAIHVLSGVWVAYERVGFSGEQYILEKGVYRNCDDWGSGNCALGSLQPVVQVGESDLH 473

Qy        172 ----MDFYSDPN 179
                  :  :| ||
Db        474 FVTKIQLFSGPN 485


RESULT 4
US-10-723-860-1946
; Sequence 1946, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
;  APPLICANT: Aziz, Natasha
;  APPLICANT:  Ginsburg, Wendy M.
;  APPLICANT:  Zlotnik, Albert
;  TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
;  TITLE OF INVENTION:  Methods for Screening for Soft Tissue Sarcoma Modulators
;  FILE REFERENCE: 05882.0193.NPUS01
;  CURRENT APPLICATION NUMBER: US/10/723,860
;  CURRENT FILING DATE:  2003-11-26
;  PRIOR APPLICATION NUMBER: 60/429,739
;  PRIOR FILING DATE: 2002-11-26
;  NUMBER OF SEQ ID NOS: 8393
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1946
;   LENGTH: 197
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-723-860-1946

  Query Match             23.2%;  Score 257;  DB 17;  Length 197;
  Best Local Similarity   34.6%;  Pred. No. 2.1e-18;
  Matches   62;  Conservative   35;  Mismatches   66;  Indels   16;  Gaps    7;

Qy          8 EDRGFQGHYYSCNSDCPNL--QPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGDYPD 65
              ::  |||       ||| |:  :    |  |::| :| |: :| |::|| |: | :||||
Db         18 DEEDFQGRRCRLLSDCANVCERGGLPRVRSVKVENGVWVAFEYPDFQGQQFILEKGDYPR 77

Qy         66 YQQWMGF----NDSIRSCR--LIPQHTGTFRMRIYERDDFRGQMSEITDDCPSLQDR-FH 118
              :  | |     :: : | |  |    |  : |: ::| |:|:|   :: || |||    :
Db         78 WSAWSGSSSHNSNQLLSFRPVLCANHNDS-RVTLFEGDNFQGCKFDLVDDYPSLPSMGWA 136

Qy        119 LTEVHSLNVLEGSWVLYEMPSYRGRQYLL----RPGEYRRYLDWG--AMNAKVGSLRRV 171
              :| || |  |:|| |: | ||| ||:|      ||:  | : |  |   :: |:|||
Db        137 SKDVGSLKVSSGAWVAYQYPGYRGYQYVLERDRHSGEFCTYGELGTQAHTGQLQSIRRV 195


RESULT 5
US-09-925-298-655
; Sequence 655, Application US/09925298
; Publication No. US20020039764A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103
;  CURRENT APPLICATION NUMBER: US/09/925,298
;  CURRENT FILING DATE:  2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 655
;   LENGTH: 97
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
```

```
;   LOCATION: (38)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (91)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-298-655

  Query Match           22.8%;  Score 253;  DB 9;  Length 97;
  Best Local Similarity  51.2%;  Pred. No. 2.3e-18;
  Matches   41;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          3 RIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGD 62
              :| | ||: ||| | |: || :  | |||||:|  | | :|||||: |: | | :|:
Db         15 QITFYEDKNFQGRRYDCDCDCADXHTYLSRCNSIKVEGGTWAVYERPNFAGYMYILPQGE 74

Qy         63 YPDYQQWMGFNDSIRSCRLI 82
              ||:||:||| || : | | :
Db         75 YPEYQRWMGLNDRLSSXRAV 94


RESULT 6
US-10-102-806-655
; Sequence 655, Application US/10102806
; Publication No. US20030054421A1
; GENERAL INFORMATION:
;  APPLICANT: Rosen et al.
;  TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
;  FILE REFERENCE: PA103P1C1
;  CURRENT APPLICATION NUMBER: US/10/102,806
;  CURRENT FILING DATE:  2002-03-22
;  PRIOR APPLICATION NUMBER: 09/925,298
;  PRIOR FILING DATE: 2001-08-10
;  PRIOR APPLICATION NUMBER: PCT/US00/05881
;  PRIOR FILING DATE: 2000-03-08
;  PRIOR APPLICATION NUMBER: 60/124,270
;  PRIOR FILING DATE: 1999-03-12
;  NUMBER OF SEQ ID NOS: 846
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 655
;   LENGTH: 97
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: SITE
;   LOCATION: (38)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
;   NAME/KEY: SITE
;   LOCATION: (91)
;   OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-102-806-655

  Query Match           22.8%;  Score 253;  DB 14;  Length 97;
  Best Local Similarity  51.2%;  Pred. No. 2.3e-18;
  Matches   41;  Conservative   14;  Mismatches   25;  Indels    0;  Gaps    0;

Qy          3 RIKFKEDRGFQGHYYSCNSDCPNLQPYFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGD 62
              :| | ||: ||| | |: || :  | |||||:|  | | :|||||: |: | | :|:
Db         15 QITFYEDKNFQGRRYDCDCDCADXHTYLSRCNSIKVEGGTWAVYERPNFAGYMYILPQGE 74

Qy         63 YPDYQQWMGFNDSIRSCRLI 82
              ||:||:||| || : | | :
Db         75 YPEYQRWMGLNDRLSSXRAV 94


RESULT 7
US-10-104-047-2339
; Sequence 2339, Application US/10104047
; Publication No. US20030236392A1
; GENERAL INFORMATION:
;  APPLICANT: HELIX RESEARCH INSTITUTE
;  TITLE OF INVENTION: No. US20030236392A1el full length cDNA
;  FILE REFERENCE: H1-A0105
;  CURRENT APPLICATION NUMBER: US/10/104,047
;  CURRENT FILING DATE:  2002-03-25
;  PRIOR APPLICATION NUMBER:
;  PRIOR FILING DATE:
;  NUMBER OF SEQ ID NOS: 4096
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2339
;   LENGTH: 511
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-104-047-2339

  Query Match           19.0%;  Score 210.5;  DB 14;  Length 511;
  Best Local Similarity  30.1%;  Pred. No. 4.9e-13;
  Matches   43;  Conservative   33;  Mismatches   66;  Indels    1;  Gaps    1;

Qy         30 FSRCNSIRVLSGCWMLYERPNYQGHQYFLRRGDYPDYQQWMGFNDSIRSCRLIPQHTGTF 89
              | |  ||||: | |: ||: :::| |: |  ||: |       :  | | | :  :
Db         54 FHRIGSIRVIGGVWVAYEKEHFKGQQFLLEEGDFEDSNACGALSSPILSFRYLQANFIES 113

Qy         90 RMRIYERDDFRGQMSEITD-DCPSLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLR 148
              : ::| |   |:  :||: :  |::   ::  |::|  | || |:   : | ||:|
Db        114 SVTLFESDLESGKFIDITNQEISDLEEIGFGSKTRSIHVKSGVWVAYQQKFFCGEQYILE 173

Qy        149 PGEYRRYLDWGAMNAKVGSLRRV 171
              |:|: : |||  |  : |:| :
Db        174 KGKYKCFFDWGGSNNIIMSIRPI 196


RESULT 8
US-10-106-698-5959
; Sequence 5959, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
;  APPLICANT: Ruben et al.
;  TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
;  FILE REFERENCE: PA005P1
;  CURRENT APPLICATION NUMBER: US/10/106,698
;  CURRENT FILING DATE:  2002-03-27
;  PRIOR APPLICATION NUMBER: PCT/US00/26524
;  PRIOR FILING DATE: 2000-09-28
;  PRIOR APPLICATION NUMBER: US 60/157,137
;  PRIOR FILING DATE: 1999-09-29
;  PRIOR APPLICATION NUMBER: US 60/163,280
;  PRIOR FILING DATE: 1999-11-03
;  NUMBER OF SEQ ID NOS: 8564
;  SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 5959
;   LENGTH: 138
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-10-106-698-5959

  Query Match           18.1%;  Score 200;  DB 14;  Length 138;
  Best Local Similarity  36.0%;  Pred. No. 1.2e-12;
  Matches   49;  Conservative   26;  Mismatches   47;  Indels   14;  Gaps    6;

Qy         49 PNYQGHQYFLRRGDYPDYQQWMGF----NDSIRSCR--LIPQHTGTFRMRIYERDDFRGQ 102
              |::|| |: | :|||| :  | |     :: : | |  |   |  : |: ::| |:|:|
Db          2 PDFQGQQFILEKGDYPRWSAWSGSSSHNSNQLLSFRPVLCANHNDS-RVTLFEGDNFQGC 60

Qy        103 MSEITDDCPSLQDR-FHLTEVHSLNVLEGSWVLYEMPSYRGRQYLL----RPGEYRRYLD 157
              :: || |||    :   :| || |  |:|| |: | ||| ||:|      ||:  | :
Db         61 KFDLVDDYPSLPSMGWASKDVGSLKVSSGAWVAYQYPGYRGYQYVLERDRHSGEFCTYGE 120

Qy        158 WG--AMNAKVGSLRRV 171
              |  |   :: |:|||
Db        121 LGTQAHTGQLQSIRRV 136


RESULT 9
US-09-864-761-44293
; Sequence 44293, Application US/09864761
; Patent No. US20020048763A1
```

```
; GENERAL INFORMATION:
;  APPLICANT: Penn, Sharron G.
;  APPLICANT:  Rank, David R.
;  APPLICANT:  Hanzel, David K.
;  APPLICANT:  Chen, Wensheng
;  TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
;  TITLE OF INVENTION:  GENE EXPRESSION ANALYSIS BY MICROARRAY
;  FILE REFERENCE: Aeomica-X-1
;  CURRENT APPLICATION NUMBER: US/09/864,761
;  CURRENT FILING DATE:  2001-05-23
;  PRIOR APPLICATION NUMBER: US 60/180,312
;  PRIOR FILING DATE: 2000-02-04
;  PRIOR APPLICATION NUMBER: US 60/207,456
;  PRIOR FILING DATE: 2000-05-26
;  PRIOR APPLICATION NUMBER: US 09/632,366
;  PRIOR FILING DATE: 2000-08-03
;  PRIOR APPLICATION NUMBER: GB 24263.6
;  PRIOR FILING DATE: 2000-10-04
;  PRIOR APPLICATION NUMBER: US 60/236,359
;  PRIOR FILING DATE: 2000-09-27
;  PRIOR APPLICATION NUMBER: PCT/US01/00666
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00667
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00664
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00669
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00665
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00668
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00663
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00662
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00661
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: PCT/US01/00670
;  PRIOR FILING DATE: 2001-01-30
;  PRIOR APPLICATION NUMBER: US 60/234,687
;  PRIOR FILING DATE: 2000-09-21
;  PRIOR APPLICATION NUMBER: US 09/608,408
;  PRIOR FILING DATE: 2000-06-30
;  PRIOR APPLICATION NUMBER: US 09/774,203
;  PRIOR FILING DATE: 2001-01-29
;  NUMBER OF SEQ ID NOS: 49117
;  SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 44293
;   LENGTH: 83
;   TYPE: PRT
;   ORGANISM: Homo sapiens
;   FEATURE:
;   OTHER INFORMATION: MAP TO AC005996.2
;   OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 0.95
;   OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 0.94
;   OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.9
;   OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL =   1
;   OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.95
;   OTHER INFORMATION: SWISSPROT HIT: P26444, EVALUE 8.00e-19
;   OTHER INFORMATION: EST_HUMAN HIT: AI142108.1, EVALUE 4.00e-14
US-09-864-761-44293

  Query Match             14.7%;  Score 163;  DB 9;  Length 83;
  Best Local Similarity   38.0%;  Pred. No. 4.4e-09;
  Matches   30;  Conservative   13;  Mismatches   34;  Indels    2;  Gaps    1;

Qy          4 IKFKEDRGFQGHYYSCNSDCPNLQP--YFSRCNSIRVLSGCWMLYERPNYQGHQYFLRRG 61
              |    | : | |        || | |    : :| ||| | || |: :  |:::| |: |  |
Db          1 ITLYEGKHFTGQKLEVFGDCDNFQDRGFMNRVNSIHVESGAWVCFNHPDFRGQQFILEHG 60

Qy         62 DYPDYQQWMGFNDSIRSCR 80
                ||||: :|   :| : |||
Db         61 DYPDFFRWNSHSDHMGSCR 79


RESULT 10
US-09-808-602-62
; Sequence 62, Application US/09808602
; Patent No. US20020155115A1
; GENERAL INFORMATION:
;  APPLICANT: Vernet, Corine A
;  APPLICANT:  Fernandes, Elma
;  APPLICANT:  Shimkets, Richard A
;  APPLICANT:  Herrman, John L
;  APPLICANT:  Majumder, Kumud
;  APPLICANT:  Mishra, Vishnu
;  APPLICANT:  Mezes, Peter S
;  APPLICANT:  MacDougall, John
;  TITLE OF INVENTION: No. US20020155115A1el Proteins and Nuclec Acids Encoding Same
;  FILE REFERENCE: 15966-697 CIP
;  CURRENT APPLICATION NUMBER: US/09/808,602
;  CURRENT FILING DATE:  2001-03-14
;  PRIOR APPLICATION NUMBER: 09/800,198
;  PRIOR FILING DATE: 2001-03-05
;  PRIOR APPLICATION NUMBER: 60/186,596
;  PRIOR FILING DATE: 2000-03-03
;  NUMBER OF SEQ ID NOS: 114
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 62
;   LENGTH: 30
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-808-602-62

  Query Match             13.3%;  Score 147;  DB 9;  Length 30;
  Best Local Similarity   90.0%;  Pred. No. 5.7e-08;
  Matches   27;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy        140 YRGRQYLLRPGEYRRYLDWGAMNAKVGSLR 169
              |||||||||| ||||||||| |||||| |
Db          1 YRGRQYLLRPGMYRRYLDWGAANAKVGSFR 30


RESULT 11
US-10-395-241-18
; Sequence 18, Application US/10395241
; Publication No. US20040038367A1
; GENERAL INFORMATION:
;  APPLICANT: YAOI, Katsuro
;  APPLICANT:  MITSUISHI, Yasushi
;  TITLE OF INVENTION: NOVEL XYLOGLUCAN OLIGOSACCHARIDE-DEGRADING ENZYME, POLYNUCLEOTIDE
;  TITLE OF INVENTION:  ENCODING THE SAME, AND METHOD OF PREPARING THE ENZYME
;  FILE REFERENCE: Q73756
;  CURRENT APPLICATION NUMBER: US/10/395,241
;  CURRENT FILING DATE:  2003-03-25
;  PRIOR APPLICATION NUMBER: JP 2002-83433
;  PRIOR FILING DATE: 2002-03-25
;  NUMBER OF SEQ ID NOS: 19
;  SOFTWARE: PatentIn version 3.2
; SEQ ID NO 18
;   LENGTH: 826
;   TYPE: PRT
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Xyloglucan Oligosaccharide-Degrading Enzyme with Histidine Tag
US-10-395-241-18

  Query Match             11.1%;  Score 123.5;  DB 15;  Length 826;
  Best Local Similarity   30.2%;  Pred. No. 0.001;
  Matches   39;  Conservative   10;  Mismatches   41;  Indels   39;  Gaps    4;

Qy         93 IYERDDFRGQMSEITDDCPSLQDRFHLTEVHSLNVLEGSWVLYEMPSYRGRQYLLRPGEY 152
              :|  ||    : : |             : |: :   |  ::   |  || ||   |
Db        713 LYRSDDNGSTWTRVND-----------QEHNYS---GPTMIEADPKVYGRVYLGTNGRG 757
```

```
Qy        153 RRYLDW-----------GAMNAKVGS-----------LRRVMDFYSDPNSSSVDKLAA 188
              | |                 | : |:               |    ||||||||||||
Db        758 IVYADLTNKKSNEEKSTAKCANGQKGTHCYVKKEIWVPWCHAVPWLISDPNSSSVDKLAA 817

Qy        189 ALEHHHHHH 197
              |||||||||
Db        818 ALEHHHHHH 826
```

RESULT 12
US-10-360-053-2
; Sequence 2, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibo
; TITLE OF INVENTION: Fragments via Antiparallel Hetergeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-2

```
  Query Match             11.1%;  Score 123;  DB 14;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 0.00016;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        176 SDPNSSSVDKLAAALEHHHHHH 197
              ||||||||||||||||||||||
Db        147 SDPNSSSVDKLAAALEHHHHHH 168
```

RESULT 13
US-10-360-053-4
; Sequence 4, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibo
; TITLE OF INVENTION: Fragments via Antiparallel Hetergeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 168
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-4

```
  Query Match             11.1%;  Score 123;  DB 14;  Length 168;
  Best Local Similarity   100.0%;  Pred. No. 0.00016;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        176 SDPNSSSVDKLAAALEHHHHHH 197
              ||||||||||||||||||||||
Db        147 SDPNSSSVDKLAAALEHHHHHH 168
```

RESULT 14
US-10-010-160-68
; Sequence 68, Application US/10010160
; Publication No. US20030103999A1
; GENERAL INFORMATION:
; APPLICANT: Rosey, Everett L.
; APPLICANT: Strugnell, Richard A.
; APPLICANT: Good, Robert T,
; APPLICANT: King, Kendall W.
; TITLE OF INVENTION: NOVEL THERAPEUTIC COMPOSITIONS FOR
; TITLE OF INVENTION: TREATING INFECTION BY LAWSONIA SPP.
; FILE REFERENCE: DAVI110.001AUS
; CURRENT APPLICATION NUMBER: US/10/010,160
; CURRENT FILING DATE: 2001-11-09
; PRIOR APPLICATION NUMBER: AU PR1381
; PRIOR FILING DATE: 2000-11-10
; PRIOR APPLICATION NUMBER: US 60/249,596
; PRIOR FILING DATE: 2000-11-17
; NUMBER OF SEQ ID NOS: 68
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 68
; LENGTH: 21
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Peptide
US-10-010-160-68

```
  Query Match             10.7%;  Score 119;  DB 14;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 3e-05;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy        177 DPNSSSVDKLAAALEHHHHHH 197
              |||||||||||||||||||||
Db          1 DPNSSSVDKLAAALEHHHHHH 21
```

RESULT 15
US-10-360-053-10
; Sequence 10, Application US/10360053
; Publication No. US20030170230A1
; GENERAL INFORMATION:
; APPLICANT: Caterer, Nigel
; APPLICANT: Uttenthal, Lars O
; APPLICANT: Nielsen, Rasmus W
; TITLE OF INVENTION: Compositions and Methods for Assembly and Stabilization of Antibo
; TITLE OF INVENTION: Fragments via Antiparallel Hetergeneous Coiled-Coil Peptide Regi
; TITLE OF INVENTION: Thereof
; FILE REFERENCE: IMX-0028
; CURRENT APPLICATION NUMBER: US/10/360,053
; CURRENT FILING DATE: 2003-02-05
; PRIOR APPLICATION NUMBER: US60/354,376
; PRIOR FILING DATE: 2002-02-05
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-360-053-10

```
  Query Match              9.1%;  Score 100.5;  DB 14;  Length 151;
  Best Local Similarity   50.0%;  Pred. No. 0.031;
  Matches   20;  Conservative    2;  Mismatches    7;  Indels   11;  Gaps    1;
```

```
Qy        158 WGAMNAKVGSLRRVMDFYSDPNSSSVDKLAAALEHHHHHH 197
              | |: ||    |:           ||||||||||||||
Db        123 WEALKAKASQRRQ----------DVDKLAAALEHHHHHH 151
```

Search completed: January 14, 2005, 12:13:02
Job time : 146 secs